GFI1 proteins orchestrate the emergence of Haematopoietic Stem Cells through recruitment of LSD1

Roshana Thambyrajah[1*], Milena Mazan[1,2*], Rahima Patel[1], Victoria Moignard[3], Monika Stefanska[1], Elli Marinopoulou[1], Yaoyong Li[4], Christophe Lancrin[1,5],Thomas Clapes[6], Tarik Möröy[7], Catherine Robin[6], Crispin Miller[4], Shaun Cowley[8], Berthold Göttgens[3], Valerie Kouskoff[9,10] & Georges Lacaud[1,10].

[1]CRUK Stem Cell Biology Group, Cancer Research UK Manchester Institute, The University of Manchester, Wilmslow road, M20 4BX, UK.

[2]Haematological Cancer Genetics, Wellcome Trust Sanger Institute, Hinxton, Cambridge, CB10 1SA, UK.

[3]Department of Haematology, Cambridge Institute for Medical Research & Wellcome Trust and MRC Stem Cell Institute, Hills Road, Cambridge CB2 0XY, UK.

[4]CRUK Computational Biology Group, Cancer Research UK Manchester Institute, The University of Manchester, Wilmslow road, M20 4BX, UK

[5]European Molecular Biology Laboratory, Mouse Biology Unit, Monterotondo, Italy

[6]Hubrecht Institute, Uppsalalaan 8, 3584 CT Utrecht, Netherlands

[7]Institut de recherches cliniques de Montréal (IRCM) and département de microbiologie et immunologie, Université de Montréal, Montréal, Canada.

[8]Department of Biochemistry, University of Leicester, Lancaster Road, Leicester, LE1 9HN, UK.

[9]CRUK Stem Cell Haematopoiesis Group, Cancer Research UK Manchester Institute, The University of Manchester, Wilmslow road, M20 4BX, UK

[10]Correspondence should be addressed to georges.lacaud@cruk.manchester.ac.uk Phone: +44 (0) 161 446 8380, Fax: +44 (0) 161 446 3109 or valerie.kouskoff@cruk.manchester.ac.uk Phone +44 (0)161 446 8381.

*Co-first author

Running title: GFI1 proteins regulate stem cell formation in the AGM

ABSTRACT

In vertebrates, the first haematopoietic stem cells (HSCs) with multi-lineage and long-term repopulating potential arise in the AGM (aorta–gonad–mesonephros) region. These HSCs are generated from a rare and transient subset of endothelial cells, called haemogenic endothelium (HE), through an endothelial-to-haematopoietic transition (EHT). Here, we establish the absolute requirement of the transcriptional repressors Growth Factor Independence 1 and 1B (GFI1 and GFI1B) in this unique trans-differentiation process. We first demonstrate that *Gfi1* expression specifically defines the rare population of HE that generates emerging HSCs. We further establish that in the absence of GFI1 proteins, HSCs and haematopoietic progenitor cells are not produced in the AGM, revealing the critical requirement for GFI1 proteins in intra-embryonic EHT. Finally, we demonstrate that GFI1 proteins recruit the chromatin modifying protein LSD1, a member of the CoREST repressive complex, to epigenetically silence the endothelial program in HE and allow the emergence of blood cells.

INTRODUCTION

The vertebrate haematopoietic system is characterised by several waves of blood progenitor generation[1]. The first two waves of blood cell emergence are initiated in the extra-embryonic yolk sac and generate mainly primitive erythrocytes then erythroid-myeloid progenitors[2-5]. In the next wave, the intra-embryonic AGM region gives rise to the first haematopoietic stem cells (HSCs) with multi-lineage and long-term repopulating potential in adult recipients[6,7]. Recent studies have provided evidence for the existence of a subset of endothelial cells capable of generating blood cells; a haemogenic endothelium[8-12] (HE). Live imaging studies have revealed how during this process of endothelial-to-haematopoietic transition (EHT), spindle-shaped endothelial cells lose their epithelial morphology to adopt the typical round shape of haematopoietic cells[10,11,13-16]. Consistent with the generation of HSCs from endothelial cells, the first progenitors with HSC activity are detected in the major arteries[17] and cells with a HSC phenotype have been found in the intra-aortic haematopoietic clusters (IAHC) generated through EHT from the ventral wall of the dorsal aorta (vDA)[13,18,19].

The molecular mechanisms controlling EHT and the specification and emergence of HSCs are still poorly understood. The transcription factor RUNX1 (AML1) is a master regulator of definitive haematopoiesis that is expressed in the entire continuum of blood development from HE to mature haematopoietic cells[20-24]. *Runx1* deletion leads to mid-gestation embryonic lethality and a complete absence of definitive haematopoietic cells[20,22,25,26] as a result of a block in the generation of blood cells from HE[11,15,27]. More recently, the two highly homologous transcriptional repressors GFI1 and GFI1B were identified as downstream targets of RUNX1 and

important players during the initial wave of haematopoietic commitment in the yolk sac[28].

In this study, we investigated the molecular mechanisms controlling the generation of HSCs during the intra-embryonic wave of haematopoiesis in the AGM and demonstrate a prominent role for the transcriptional repressors GFI1 and GFI1B in this process. By combining mouse reporter lines and functional assays, we established that *Gfi1* expression is initiated in a rare population of HE, present in the vDA during HSC emergence, whereas GFI1B expression is up-regulated in emerging IAHC. Loss of GFI1 proteins impaired IAHC formation, indicating the pivotal role of these repressors during the emergence of haematopoietic cells within the AGM. Furthermore, we established that the critical function of the GFI1 proteins is mediated by the recruitment of LSD1, a member of the repressive CoREST complex. Deficiency in LSD1 prevents the morphological changes and loss of endothelial markers associated with the EHT. Accordingly, we observed that GFI1 and GFI1B directly bind to the regulatory regions of a large set of genes whose expression is maintained upon LSD1 inhibition. This set of genes is closely associated with cardiovascular development, as well as maintenance and remodelling of blood vessels. Altogether, our results demonstrate that, in the AGM, the sequential expression of GFI1 proteins orchestrate the generation of HSCs by recruiting the CoREST complex to silence the endothelial programme.

RESULTS

GFI1$^+$/GFI1B$^-$ and GFI1$^+$/GFI1B$^+$ cells mark distinct populations within the AGM

To determine the specific expression and potential functions of GFI1 and GFI1B in the emergence of HSCs, we first analysed the spatiotemporal expression pattern of GFI1 and GFI1B in the AGM. For this, a *Gfi1:H2B-Tomato* mouse reporter line (referred to as *Gfi1^{tomato}* hereafter) was generated, validated (Supplementary Fig. 1a) and crossed with the *Gfi1b$^{+/GFP}$* knock-in line[29]. The expression of both *Gfi1* and *Gfi1b* was investigated in double transgenic embryos by fluorescence activated cell sorting (FACS) and immunohistochemistry (IHC). At E9.5, when the dorsal aorta is developing within the para-aortic splanchno-pleura (P-Sp) region, *Gfi1* expression was not detected by IHC in sections of the dual reporter embryos. *Gfi1b* expression was restricted to circulating cells likely representing primitive erythroid cells (Fig. 1ai).

At E10.5, GFI1$^+$/GFI1B$^-$ cells were found either embedded within the endothelium (CD31 or VE-cadherin/CDH5) (arrowheads, Fig. 1aii, Supplementary Fig. 1b), and/or at the base of IAHCs (arrows Fig. 1aii). Some GFI1$^+$ cells in IAHCs were also positive for GFI1B (Fig. 1aii). We next analysed the expression of *Gfi1* and *Gfi1b* in the context of previously characterized markers of HE and IAHC in the AGM: RUNX1, c-KIT, CD41 and CD45. Staining for β-Galactosidase on *Runx1:nls-LacZ*/*Gfi1$^{GFP/+}$* E10.5 AGMs showed that *Gfi1* expression was restricted to a subset of RUNX1$^+$ cells within the endothelial layer (white arrows, Fig. 1b) and did not extend to the sub-aortic mesenchyme and endothelium (yellow arrows and arrowheads, Fig. 1b). Previous studies have shown that CDH5$^+$ cells in the vDA

generate IAHCs that express the blood markers c-KIT and CD41 before acquiring CD45. These $CDH5^+/CD41^+$ and $CDH5^+/CD45^+$ intermediates represent pre-HSCs that mature into directly transplantable HSCs[30-32]. We observed that $GFI1^+$ cells in the clusters co-stained for c-KIT (arrow, Fig. 1aiii), CD41 (arrow, Fig. 1aiv) and CD45 (arrow, Fig. 1av) at E10.5. Similarly at E11.5, $GFI1^+GFI1B^-$ and $GFI1^+GFI1B^+$ cells were still closely associated with endothelial lining and IAHCs at E11.5 (Fig. 2a).

In agreement with the IHC data, only a very low frequency of $GFI1^+/GFI1B^-$ or $GFI1^+/GFI1B^+$ cells were found by FACS in E10.5 and E11.5 AGMs (Fig. 2b). A large fraction of $GFI1^+/GFI1B^-$ cells were in the $CDH5^+/CD45^-$ endothelial subset or in the $CDH5^+/CD45^+$ double positive cell population (Fig. 2c). In contrast, $GFI1^+/GFI1B^+$ cells were mostly associated with the more mature $CDH5^+/CD45^+$ or $CDH5^-/CD45^+$ populations (Fig. 2c). Within the $CDH5^+$ endothelial compartment, the frequency of $GFI1^+/GFI1B^-$ was similar at E10.5 and E11.5, whereas $GFI1^+/GFI1B^+$ tended to increase (Supplementary Fig. 1c). Within this compartment, we also found a small population of cells that only expressed *Gfi1b* at E11.5 (Supplementary Fig.1d). Finally, in order to visualise the acquisition of *Gfi1b* expression by $GFI1^+$ cells, we performed *in vitro* culture of FACS sorted $CDH5^+/GFI1^+/GFI1B^-$ cells and observed the emergence of GFI1B expressing cells (Supplementary Fig. 2). Similarly, time-lapse imaging of $Gfi1^{tomato}Gfi1b^{gfp}$ embryonic sections showed the acquisition of *Gfi1b* expression by $CDH5^+/GFI1^+/GFI1B^-$ cells (Fig. 2d, Supplementary movie1).

Overall, these data revealed a dynamic and distinct expression pattern for *Gfi1* and *Gfi1b*. *Gfi1* expression is initiated at E10.5 in $RUNX1^+$ endothelial cells within the vDA and is sustained in the first emerging cells within the IAHCs, suggesting that *Gfi1* expression marks HE undergoing the transition to

haematopoiesis. *Gfi1b* expressing cells, in contrast, are mainly localized in IAHCs as c-KIT$^+$ cells and coincide with further commitment to CD41 and CD45 expression.

GFI1$^+$/GFI1B$^-$ cells mark HE while HSC activity is mostly associated with GFI1$^+$/GFI1B$^+$ expression

To demonstrate that the endothelial GFI1$^+$/GFI1B$^-$ cells correspond to HE that subsequently become GFI1$^+$/GFI1B$^+$ blood precursors, we analysed the expression of a set of 70 endothelial and haematopoietic genes in single cells of the different CDH5$^+$ populations defined by *Gfi1* and *Gfi1b* expression. These included endothelial CDH5$^+$ cells, the CDH5$^+$ cells expressing *Gfi1* found in the endothelial lining (CDH5$^+$/GFI1$^+$/GFI1B$^-$/c-KIT$^-$), or emerging in the lumen (CDH5$^+$/GFI1$^+$/GFI1B$^-$/c-KIT$^+$), the CDH5$^+$/GFI1$^+$/GFI1B$^+$ cells present in the clusters and the rare population of CDH5$^+$/GFI1$^-$/GFI1B$^+$ cells (Fig. 3a and Supplementary Fig. 3). Consistent with a developmental continuum between the three GFI1$^+$ populations, principal component analyses indicated that CDH5$^+$/GFI1$^+$/GFI1B$^-$/c-KIT$^-$ and CDH5$^+$/GFI1$^+$/GFI1B$^+$ cells formed separate clusters but were linked by the CDH5/GFI1$^+$/GFI1B$^-$/c-KIT$^+$ population (Fig. 3b). In contrast, endothelial CDH5$^+$ or CDH5$^+$/GFI1$^-$/GFI1B$^+$ cells clustered as separate entities with the latter displaying an expression profile consistent with megakaryocytes. *Gfi1* expression closely co-coincided with the expression of endothelial genes (*Dll4*, *Procr*, *Kdr*, *Egfl7*, *Notch1*) whereas the expression of *Gfi1b* was linked to haematopoietic genes such as *c-kit*, *Itga2b*, *CD45*, *Mpo* and *c-myb* (Fig. 3b). Together, these analyses suggest that CDH5$^+$/GFI1$^+$GFI1B$^-$/c-KIT$^-$, CDH5$^+$/GFI1$^+$/GFI1B$^-$/c-KIT$^+$ and CDH5$^+$/GFI1$^+$/GFI1B$^+$

are distinct but sequentially related cell populations corresponding to haemogenic endothelium, cells undergoing EHT and emerging haematopoietic cells, respectively.

By definition, HE cells require a further maturation step to generate haematopoietic progenitors. To functionally evaluate the HE potential of the GFI1$^+$ fraction, we isolated CDH5$^+$/GFI1$^-$ and CDH5$^+$/GFI1$^+$ cells from E10.5 AGM while excluding blood precursors (CD41$^+$ and/or CD45$^+$). Both populations were plated into haematopoietic assay, either directly, or following a 4-day maturation step of co-culture with OP9 cells. Direct re-plating into haematopoietic assay yielded very few colonies (Fig. 3d) in contrast to control sorted GFI1$^+$ and/or GFI1B$^+$ cells containing CD41$^+$ and CD45$^+$ cells (Supplementary Fig. 4a). Following maturation, the CDH5$^+$/GFI1$^-$/CD41$^-$/CD45$^-$ fraction did not gain any haematopoietic potential, while the CDH5$^+$/GFI1$^+$/CD41$^-$/CD45$^-$ cells generated haematopoietic colonies at a high frequency, indicating that at E10.5, HE cells were present within the CDH5$^+$/GFI1$^+$ population (Fig. 3d).

Finally, we tested whether GFI1$^+$ cells in the vDA contain HSCs. Single cell suspension of *Gfi1*$^{tomato/+}$ E11.5 AGMs were stained for CDH5 and separated into CDH5$^+$/GFI1$^+$ (that are either expressing *Gfi1b* or not) and CDH5$^+$/GFI1$^-$ cells. Both fractions were transplanted into lethally irradiated mice. Long-term primary and secondary contributions were mainly detected in recipients transplanted with the CDH5$^+$/GFI1$^+$ fraction (Fig. 3ei, Fig. Supplementary Fig. 4b and 4c), indicating that HSC activity is restricted to the CDH5$^+$/GFI1$^+$ population at E11.5. We further separated the CDH5$^+$/GFI1$^+$ cell population into CDH5$^+$/GFI1$^+$/GFI1B$^-$ and CDH5/GFI1$^+$/GFI1B$^+$ cells in order to test if there were significant differences between these two populations. Consistent with the less mature state of the

CDH5^{+}/GFI1^{+}/GFI1B^{-} cells, the repopulation potential of this fraction was lower than that of CDH5^{+}/GFI1^{+}/GFI1B^{+} cells (Fig 3eii).

GFI1 and GFI1B are required for EHT in the AGM

In order to investigate if GFI1 proteins have a functional importance in HSC emergence, we examined the presence of IAHC in E10.5 embryos. IAHC were readily observed in control het/het and knock-out embryos for either GFI1 or GFI1B (Fig. 4ai-iii and Supplementary Fig. 5a). In contrast, no c-KIT^{+} IAHCs were found in the double *Gfi1s* ko (*Gfi1*$^{GFP/GFP}$*/Gfi1b*$^{GFP/GFP}$) AGMs (Fig. 4aiv). Instead, we detected GFP^{+} cells, which were fully embedded within the lining of the vDA, suggesting a critical impairment in the EHT and IAHC generation. The presence of clusters in *Gfi1* deficient embryos might be explained by the upregulation of *Gfi1b* that could result in a functional compensation for the loss of GFI1. Indeed it has been previously demonstrated that GFI1B could functionally replace and rescue the haematopoietic defects observed in *Gfi1* deficient animals[28,33]. Accordingly, we observed the appearance of *Gfi1b* transcripts in the endothelial lining of the dorsal aorta of *Gfi1* deficient embryos by *in situ* staining (Supplementary Fig. 5b). By q-PCR, we also detected an upregulation of *Gfi1b* expression in the whole AGM lysate and more specifically in HE cells (CDH5^{+}/GFP^{+}/c-KIT^{-}*)* of *Gfi1* deficient embryos (*Gfi1*^{ko}*/Gfi1b*^{het})(Supplementary Fig. 5c). Altogether, these data suggest that GFI1 is critical for the initiation of the EHT, but that its absence can be compensated by induction of *Gfi1b* expression.

Finally, to test whether *Gfi1*^{ko}*/Gfi1b*^{ko} AGM cells possessed any

haematopoietic potential, we plated single cell suspensions of *Gfi1*^{ko}/*Gfi1b*^{ko} AGM in haematopoietic assays either directly, or after a maturation step on OP9 cells (Fig. 4b and 4c). None of the conditions led to haematopoietic colony formation, indicating an absolute requirement for GFI1 and GFI1B in the generation, and/or the subsequent maturation, of blood precursors in the AGM.

LSD1 is essential for the EHT

Both GFI1 and GFI1B have been shown to epigenetically repress transcription in MEL (murine erythroleukemia) cells by recruiting the chromatin regulatory CoREST complex[34]. The CoREST complex includes the histone demethylase LSD1 (KDM1A) and the histone deacetylases, HDAC1 and HDAC2. Upon recruitment to target sites, HDACs and LSD1 remove the activating acetylation and methylation marks from histones within their proximity[34]. Stable repression of the target loci is then further induced by the recruitment of G9 or SUV39H1, and subsequent methylation of repressive sites.

To investigate if the EHT is regulated by a similar epigenetic mechanism, we examined the consequences of pharmacological LSD1 inactivation[35] on blood emergence during the *in vitro* differentiation of ES cells. In this system, FLK1^{+} mesodermal haemangioblasts plated in methylcellulose, form tight HE cores from which round haematopoietic cells emerge to produce blast colonies[11,36]. We first tested the LSD1 inhibitor in this blast colony assay and observed that increasing amounts of the inhibitor dramatically decreased the number of blast colonies. In parallel, we detected a directly correlated increase in the number of cores,

suggesting that the LSD1 inhibition blocks the emergence of round blood cells from HE cores (Fig. 5a). We next assessed the impact of LSD1 inhibition in monolayer cultures by time-lapse imaging and FACS. In these cultures, HE cells aggregate as tight and adherent clusters that are characterised by the co-expression of endothelial (CDH5 and TIE2/TEK) and haematopoietic markers (c-KIT and CD41). These HE clusters then generate free-floating haematopoietic cells concomitantly with progressive loss of endothelial markers[9]. We observed that control cells transited through a HE stage and generated round blood cells (Supplementary movie 2). Strikingly, cultures treated with the LSD1 inhibitor generated HE clusters, but did not progress to produce free-floating blood cells (Supplementary movie 3). FACS analysis of the control cultures confirmed the presence of HE cells ($CDH5^+/CD41^+$ or $TIE2^+/CD41^+$) and of the subsequent $CD41^+$ single positive blood cells (Fig. 5b). In contrast, in the presence of the inhibitor, very few $CD41^+$ single positive blood cells were detected, suggesting a block in EHT. We confirmed this striking phenotype in ES cells carrying floxed *Lsd1* alleles[37]. Upon induction of deletion, we observed a similar defect in the generation of round blood cells (Supplementary movie 4 and 5, and Supplementary Fig. 6a). Finally, consistent with previous reports[38,39], we observed a decrease in cell proliferation and an increase in apoptosis upon *Lsd1* deletion during *in vitro* differentiation (Supplementary Fig. 6b and 6c).

To correlate this observation with the activity of GFI1 proteins, we assessed the *in vitro* phenotype resulting from the combined loss of *Gfi1* and *Gfi1b*. As observed for LSD1 deficiency, differentiating $Gfi1^{ko}/Gfi1b^{ko}$ ES cells formed HE clusters, but the emergence of round haematopoietic cells (Fig. 5c), as well as the down-regulation of CDH5 and TIE2 were not observed (Fig. 5d). To investigate if

LSD1 activity was also required in the AGM for the generation of blood cells, we analysed *Lsd1* expression and performed *ex-vivo* LSD1 inhibition in E9.5 P-Sp explants and E10.5 embryonic sections of *Gfi1*$^{tomato/+}$/*Gfi1b*$^{GFP/+}$ embryos. We observed that *Lsd1* is ubiquitously expressed in the mouse embryo, consistent with previous report[37] (Supplementary Fig. 7a). Following culture of control E9.5 P-Sp explants, GFP$^+$ (*Gfi1b* expressing) cells were found to be negative for the endothelial markers CD31 or CDH5 by FACS and were detected as clusters of cells by confocal microscopy (Fig. 5e, Supplementary movie 6). In contrast, upon LSD1 inhibition, part of GFI1B$^+$ cells were still CDH5$^+$ (Fig. 5e). In addition, some *Gfi1b* expressing cells were still found embedded within the CD31$^+$ endothelial lining of the vDA (Supplementary movie 7) but expressed the haematopoietic marker CD45 (Fig. 5eii and Supplementary Fig. 7b). Similarly, in culture of E10.5 sections treated with the LSD1 inhibitor, *Gfi1* and *Gfi1b* expressing cells remained embedded in the endothelial lining of the vDA (Fig. 5f and Supplementary Fig. 7c). Altogether, these results indicate the critical requirement of LSD1 for the emergence of blood cells, both in differentiating ES cells and in the AGM.

GFI1 and GFI1B shut down the HE programme by recruiting LSD1 to target genes

In order to identify the genome-wide molecular changes induced by GFI1 and GFI1B through the recruitment of LSD1 during EHT, we compared global gene expression profiles of LSD1-inhibited ES-derived HE cells (CDH5$^+$/CD41$^+$) and the corresponding CD41 compartment of control cells (CDH5$^+$/CD41$^+$ HE and CD41$^+$ haematopoietic cells). We found 244 protein-coding transcripts present at higher,

and 165 at lower, levels in LSD1-inhibited cells (Fig. 6a). Transcripts with higher levels in LSD1-inhibited cells were enriched for genes implicated in cardiovascular system development, whereas transcripts with lower levels of expression were associated with haematological system development/function and cell morphology (Fig. 6b). In parallel, we also mapped GFI1 and GFI1B binding sites in ES-derived HE cells using the DamID (DNA adenine methyltransferase Identification) strategy[40,41] (Fig. 6c, Supplementary Fig. 7d). Consistent with their potential function in HE, the genes bound by GFI1 and/or GFI1B were highly enriched for transcripts involved in integrin, actin cytoskeleton, and gap junction signalling (Fig. 6d); pathways that need to be modulated for EHT. Next, we overlapped the lists of genes bound by GFI1 and/or GFI1B with the list of genes expressed at higher levels when LSD1 activity is blocked. The resulting list of 78 candidate genes (Fig. 7a) contained genes with well-established roles in stem cells (*Lgr5, Lin28A, Sall1*), as well as in cardiovascular development, blood vessels maintenance and remodelling (*Egfl7, Nos3, Pcdh12, Ptprb, Sox7, Cdh5, Esam, Hey2, Hand2, ID1, ID3*). Accordingly, this list was significantly enriched for genes involved in development and, in particular, in cardio-vascular system development (Fig. 7b). Interestingly, 33 out of the 78 genes have been previously shown to be bound by RUNX1 in ChIP-seq experiments during EHT[42] (Fig. 7c and genes highlighted in red in Fig. 7a). *In silico* analysis also indicated a frequent occurrence of RBPJ binding motifs within, or close to the GFI1 and GFI1B peaks (Fig. 7d and genes indicated by an asterisk in Fig. 7a), suggesting that both RUNX1 and *Notch* signalling could additionally regulate the transcription of these genes. Finally, the *S1P* signalling pathway, implicated in vascular permeability[43], was found to be significantly associated with genes present in this list (Fig. 7b, bottom panel). Altogether, our molecular data strongly suggests that during

EHT, GFI1 and GFI1B bind to genes associated with the global maintenance of endothelial identity, and epigenetically silences this programme through the recruitment of the CoREST complex to orchestrate the generation of HSCs and blood progenitors (Fig. 7e).

DISCUSSION

Recent studies have conclusively demonstrated the existence of HE cells that have the unique ability to give rise to HSCs and blood precursors. However, the precise nature of this transient and rare subpopulation of endothelial cells, and how they generate different types of blood cells via EHT, remains poorly understood.

In this study, we revealed an unrecognized requirement for the two transcriptional repressors GFI1 and GFI1B for HSC emergence in the AGM. Indeed, we did not detect any IAHCs or any haematopoietic potential in E10.5 *Gfi1*^{ko}/*Gfi1b*^{ko} AGMs, even following further maturation, indicating an absolute requirement for GFI1 and GFI1B for the commitment and generation of haematopoietic cells. Interestingly this developmental block is not observed in *Gfi1* single KO suggesting that the critical function of GFI1 in HE cells could be rescued by the compensatory expression of *Gfi1b* that we observed in these cells. In the double KO embryos, discrete GFP^{+} cells were observed embedded within the endothelium of the vDA, indicating that HE cells are specified in the absence of GFI1 and GFI1B proteins, but are unable to progress through EHT. This is reminiscent of the situation in the yolk sac where the double KO GFP^{+} cells remain clustered together and are not disseminated in the vasculature[28]. In contrast, the absence of blood potential in *Gfi1*^{ko}/*Gfi1b*^{ko} AGMs is distinct from the situation in the yolk sac where GFP^{+} cells were shown to display haematopoietic potential when dissociated and replated[28]. These results suggest differences in the molecular regulation during the generation of blood cells in the yolk sac and in the AGM. This could also reflect the existence of different types of HE at these two sites, as previously suggested[44]. Interestingly, this

dichotomy mirrors the different requirements for Notch signalling which is dispensable for yolk sac haematopoiesis but essential for AGM haematopoiesis[45-48].

Detailed analysis of the EHT has so far been challenging given the limited number of genetic markers specifically defining HE. Our results suggest that *Gfi1* expression identifies the specific subset of cells in the endothelial lining that are actively undergoing the transition to the haematopoietic fate. As such, GFI1 delineates a more refined cell population than previously described markers such as the endothelial markers SOX17[49] and SOX7[50], that are widely expressed in endothelial cells, or the haematopoietic marker SCA-1[51] or RUNX1[20,52] that are extensively expressed in other AGM cell populations.

We also revealed here, that LSD1, a member of the CoREST repressive complex, is critical for the EHT process. *Lsd1* deficiency phenocopies the developmental block observed in *Gfi1*^{ko}/*Gfi1b*^{ko}, strongly suggesting that the repressive function of the *Gfi1*s during EHT is mediated by the CoREST complex. Our molecular analysis identified a specific set of 78 genes that are bound and repressed by GFI1 and GFI1B during EHT. This list not only contains genes, that have not been previously identified as GFI or GFI1B targets[53,54], but also some genes that have known, or suspected, functions in stem cell and endothelial cells. This list also harbours a wealth of new candidates potentially implicated in the EHT. In addition, a large proportion of these genes has been shown to be bound by RUNX1[42] and contains RBPJ binding motif suggesting that GFI1 repression, RUNX1 binding and Notch signalling overlap and cooperate to orchestrate the EHT process in the AGM.

Altogether, our study provides mechanistic insights and valuable novel candidates and leads for future studies of the molecular programme governing HE specification and the generation of blood cells and HSCs in the AGM. Deciphering the regulatory pathways that specifically control this EHT process is essential for precisely recapitulating this programme during ES cell differentiation[55-58] or during reprogramming[59-61] to ultimately generate HSCs *in vitro*.

Figure Legends

Figure 1 GFI1 and GFI1B define distinct cell populations in the E10.5 AGM

(a i-v) IHC on E9.5 and E10.5 *Gfi1*^{tomato}/ *Gfi1b*^{GFP} embryos for CD31 and CDH5 (both red), GFI1 (cyan), GFI1B (green) and c-KIT, CD41 and CD45 (yellow). (**i**) E 9.5 P-Sp. (**ii**) GFI1^{+}/GFI1B^{-} cells (arrowhead) and GFI1^{+}/GFI1B^{+} cells (arrow) are indicated. (**iii**) GFI1 and c-KIT are co-expressed (white arrow). (**iv**) CD41 and GFI1 are co-expressed (white arrow). (**v**) CD45, GFI1 and GFI1B co-expression (white arrow). (**b**) Staining for Runx1 (beta-Galactosidase) (yellow), CD31 (red) and GFI1 (cyan). Co-expression of CD31, GFI1 and RUNX1 (white arrow), CD31 and RUNX1 (yellow arrow) or only RUNX1 expression (yellow arrowhead) are indicated. Representative images of AGM sections derived from more than 5 embryos from different litters are shown. Scale bars =10 um.

Figure 2 GFI1 marks haemogenic endothelial cells in the E10.5 AGM

(**ai** and **aii**) IHC for CDH5 (red), GFI1 (cyan), GFI1B (green) and CD45 (yellow) at E11.5. (**i**) GFI1^{+}/GFI1B^{-} (arrowhead) and GFI1^{+}/GFI1B^{+} cells (arrow) are indicated. (**ii**) GFI1^{+}/GFI1B^{+} cells co-expressing CD45 (arrow) are indicated (Representative images of AGM sections derived from more than 3 embryos from different litters are shown). (**b**) FACS analysis of Gfi1 (TOMATO) and Gfi1b (GFP) expression in E10.5 and E11.5 AGM (Representative FACS data from more than 2 different litters are shown). (**c**) FACS analysis of CDH5 and CD45 expression by GFI1^{+}/GFI1B^{-} and

GFI1^{+}/GFI1B^{+} cells in E10.5 and E11.5 AGMs (Representative FACS data from more than 2 different litters are shown). (**d**) Images from time-lapses performed on *Gfi1^{tomato}Gfi1b^{gfp}* E10.5 sections stained with CD31 (n= 3 independent experiments with AGMs from 2 or more embryos, one representative sequence of imaging is shown). Scale bar = 10 μm.

Figure 3 GFI1^{+} cells acquire GFI1B expression and habour HSC activity

(**a**) Schematic representation of the sorting strategy for the single cell PCRs. (**b**) Principal component analysis for cells and genes analysed (left panel). Prinicipal component analysis for endothelial and haematopoietic genes expressed by CDH5^{+}/GFI1^{+}/c-KIT^{-} and CDH5^{+}/GFI1^{+}/GFI1B^{+} cells (right panel). (**c**) Violin plots of single cell PCR data for the indicated genes (data is derived from one single Fluidigm single cell PCR run performed with n= 96 AGM cells from 3 different litters). (**d**) Haemogenic potential of E10.5 CDH5^{+}/GFI1^{+}/CD41^{-}/CD45^{-} and CDH5^{+}/GFI1^{-}/CD41^{-}/CD45^{-} cells. The CDH5^{+}/GFI1^{+}/CD41^{-}/CD45^{-} and CDH5^{+}/GFI1^{-}/CD41^{-}/CD45^{-} populations were either directly re-plated or replated after a co-aggregation step with OP-9 cells for 4 days (2 independent experiments were performed with AGMs from 2 and 9 embryos. Data from one representative experiment is shown; the source data for the second experiment can be found in Statistics source data). (**e**) (**i**) Long term repopulation potential of E11.5 CDH5^{+}/GFI1^{+} and CDH5^{+}/GFI1^{-} cells from *Gfi1^{tomato}* AGMs. Donor contributions to the LSK compartment at 17 weeks after transplantation are shown (one independent experiment with AGM cells sorted from one litter). (**ii**) E11.5 *Gfi1^{tomato}Gfi1b^{gfp}* AGMs were dissociated and the CDH5^{+}/GFI1^{+}/GFI1B^{-} and CDH5^{+}/GFI1^{+}/GFI1B^{+} cells were injected intravenously

into irradiated NSG mice. Donor contributions to the LSK compartment at 16 weeks after transplantation are shown (one independent experiment with AGM cells sorted from one litter).

Figure 4 GFI1 and GFI1B are essential for intra-aortic Cluster formation

(**a**) IHC on *Gfi1*^{het}/*Gfi1b*^{het} (*Gfi1*$^{GFP/+}$/*Gfi1b*$^{GFP/+}$), *Gfi1*^{KO}/*Gfi1b*^{het} (*Gfi1*$^{GFP/GFP}$/*Gfi1b*$^{GFP/+}$), *Gfi1*^{het}/*Gfi1b*^{KO} (*Gfi1*$^{GFP/+}$/*Gfi1b*$^{GFP/GFP}$) and *Gfi1*^{KO}/*Gfi1b*^{KO} (*Gfi1*$^{GFP/GFP}$/*Gfi1b*$^{GFP/GFP}$) E10.5 AGMs for CD31 (red), GFI1 and GFI1B (green) and c-KIT (yellow) (Representative images of AGM sections derived from more than 4 embryos from different litters are shown). (**b** and **c**) Haematopoietic potential of *Gfi1*^{het}/*Gfi1b*^{het} and *Gfi1*^{ko}/*Gfi1b*^{ko} AGMs. Cells were re-plated either directly (**b**) or after a co-aggregation step with OP-9 cells for 4 days (one independent experiment with AGM cells sorted from one litter) (**c**) into haematopoietic colony assays (one independent experiment with AGM cells sorted from one litter). Scale bar = 10um.

Figure 5 LSD1 is critical for EHT

(**a**) ES-derived FLK1^{+} cells were either cultured with DMSO or different concentrations of the LSD1 inhibitor in blast assays (one representative experiment out of 2 independent experiments is shown, data from the second experiment can be found in statistical source data). (**b**,**d**) CD41^{+} single (black arrow), TIE-2^{+}/CD41^{+} and CDH5^{+}/CD41^{+} double positive (grey arrow) cells are indicated. (**b**) FACS on day 3 monolayer cultures of ES-derived FLK1^{+} cells treated with DMSO or LSD1 inhibitor

(300nM) (Representative FACS plots of more than 6 independent experiments). (**c**) Time-lapse images of monolayer cultures (Day 3) of *Gfi1*^{ko}/*Gfi1b*^{ko} or *Gfi1*^{het}/*Gfi1b*^{het} ES-derived FLK1^{+} cells (Representative images of 2 independent experiments). Scale bar: 50um. (**d**) FACS on cells isolated from day 5 EBs of *Gfi1*^{ko}/*Gfi1b*^{ko} or *Gfi1*^{het}/*Gfi1b*^{het} (Representative FACS plots of more than 6 independent experiments). (**e**) E9.5 P-Sp *Gfi1*^{tomato}/ *Gfi1b*^{GFP} explant cultures treated with DMSO (Ctrl) or LSD1 inhibitor (300-500nM). (**i**) The explants were either analysed by FACS for the presence of GFI1B^{+} cells in the CDH5^{+} endothelial compartment (Representative data of more than 3 independent experiments) or (**ii**) by imaging to generate 3D reconstruction of E9.5 AGM explants cultured with DMSO (DMSO ctrl) or 500nM of the LSD1 inhibitor (LSD1 inhib) (Representative data of 3 independent experiments). IHC for CD31 (red), GFI1 (cyan) and GFI1B (green) was performed prior to imaging. (**iii**) Imaging of E10.5 *Gfi1*^{tomato}/ *Gfi1b*^{GFP} AGM sections after overnight *ex vivo* culture with DMSO (DMSO ctrl) or 500nM of the LSD1 inhibitor (LSD1 inhib) (Representative data of 3 independent experiments). Scale bars: 20μm.

Figure 6 Molecular programme governed by LSD1 and GFI1/1B

(a) Heat map of differentially expressed genes between ES-derived CD41^{+} cells Day 2 liquid blast culture in presence of LSD1 inhibitor (300nM) or DMSO. (**b**) Molecular function of the genes expressed at higher, and lower levels, in the LSD1 inhibited cells compared to control cells. (**c**) Schematic representation of the DamID technique. (**d**) GO terms of genes bound by GFI1 and GFI1B in day 2 Liquid Blast derived HE cells.

Figure 7 Identification of genes repressed by GFI1 and GF1B in HE

(**a**) Overlap of genes bound by *Gfi1* and *Gfi1b* and expressed at higher levels of transcription upon LSD1 inhibition. Genes previously shown to harbour RUNX1 ChIP-seq peaks are indicated in red and genes with RBPJ binding motifs within 100bp of GFI1 or GFI1B peaks are indicated by an asterisk. (**b**) Functional classification and pathway enrichment analysis for the 78 candidate genes. (**c**) Pie chart representing the number of genes with/out RUNX1 Chip-seq binding peaks. (**d**) RBPJ binding motif identified by HOMER in 100bp proximity of GFI1 and/or GFI1B peaks. (**e**) Model of the expression patterns and function of GFI1 and GFI1B during HE specification and EHT in the AGM.

Acknowledgements

We thank the staff at the Advanced Imaging, animal facility, Molecular Biology Core facilities and Flow Cytometry of CRUK Manchester Institute for technical support and Michael Lie-A-Ling and Elli Marinopoulou for initiating the DamID-PIP bioinformatics project. We thank members of the Stem Cell Biology group, the Stem Cell Haematopoiesis groups and Martin Gering for valuable advice and critical reading of the manuscript. Work in our laboratory is supported by the Leukaemia and Lymphoma Research Foundation (LLR), Cancer Research UK (CRUK) and the Biotechnology and Biological Sciences Research Council (BBSRC). SC is the recipient of an MRC senior fellowship (MR/J009202/1).

Author Contributions

RT designed and performed most of the experiments, analyzed the data and wrote the manuscript, MM initiated the project, designed, performed experiments and analyzed the data. RP, VM, MS and CL designed and performed experiments. EM and YL performed bioinformatics analysis on the sequencing data and microarray. TC, TM, CR, CM, SC and BG contributed valuable tools and protocols. VK and GL designed and supervised the research project, analyzed the data, and wrote the manuscript.

Disclosure of Conflicts of Interest

The authors declare no competing financial interests.

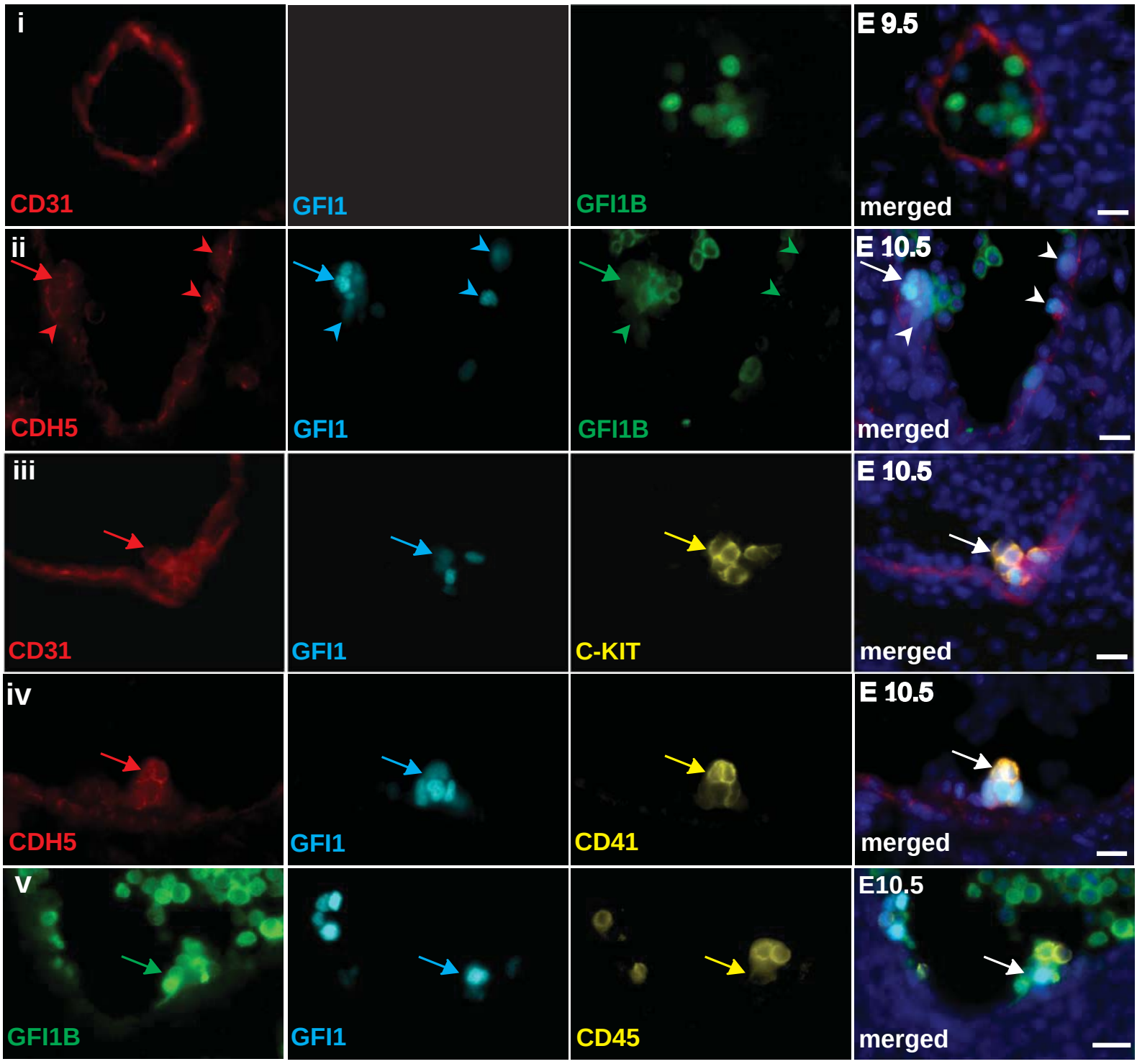
References

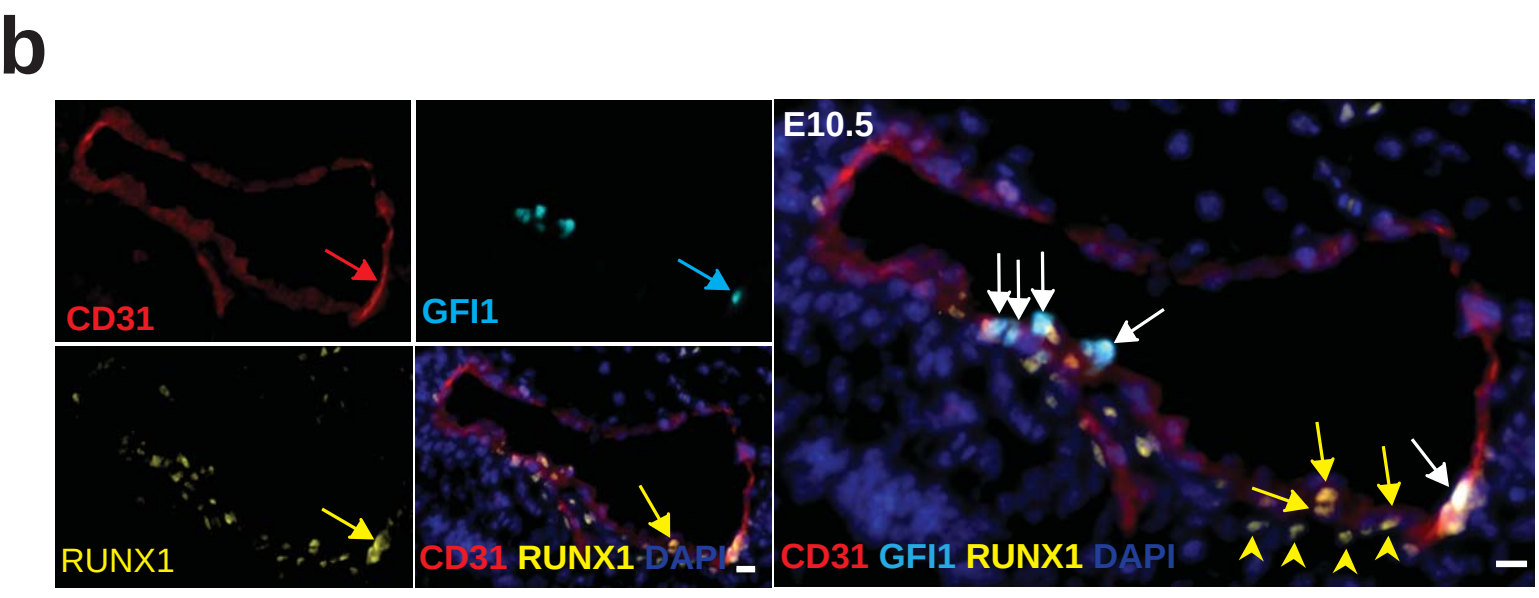
1. Costa, G., Kouskoff, V. & Lacaud, G. Origin of blood cells and HSC production in the embryo. *Trends in Immunology* **33,** 215–223 (2012).
2. Palis, J., Robertson, S., Kennedy, M., Wall, C. & Keller, G. M. Development of erythroid and myeloid progenitors in the yolk sac and embryo proper of the mouse. *Development* **126,** 5073–5084 (1999).
3. Frame, J. M., McGrath, K. E. & Palis, J. Erythro-myeloid progenitors: 'definitive' hematopoiesis in the conceptus prior to the emergence of hematopoietic stem cells. *Blood Cells Mol Dis* **51,** 220–225 (2013).
4. McGrath, K. E. *et al.* Distinct Sources of Hematopoietic Progenitors Emerge before HSCs and Provide Functional Blood Cells in the Mammalian Embryo. *CellReports* **11,** 1892–1904 (2015).
5. Frame, J. M., Fegan, K. H., Conway, S. J., McGrath, K. E. & Palis, J. Definitive Hematopoiesis In The Yolk Sac Emerges from Wnt-responsive Hemogenic Endothelium Independently Of Circulation and Arterial Identity. *Stem Cells* (2015). doi:10.1002/stem.2213
6. Medvinsky, A. J. & Dzierzak, E. A. Definitive hematopoiesis is autonomously initiated by the AGM region. *Cell* **86,** 897–906 (1996).
7. Müller, A. M., Medvinsky, A. J., Strouboulis, J., Grosveld, F. & Dzierzak, E. A. Development of hematopoietic stem cell activity in the mouse embryo. *Immunity* **1,** 291–301 (1994).
8. Jaffredo, T., Gautier, R., Eichmann, A. & Dieterlen-Lièvre, F. Intraaortic hemopoietic cells are derived from endothelial cells during ontogeny. *Development* **125,** 4575–4583 (1998).
9. Zovein, A. C. *et al.* Fate tracing reveals the endothelial origin of hematopoietic stem cells. *Cell Stem Cell* **3,** 625–636 (2008).
10. Eilken, H. M., Nishikawa, S., Nishikawa, S. & Schroeder, T. Continuous single-cell imaging of blood generation from haemogenic endothelium. *Nature* **457,** 896–900 (2009).
11. Lancrin, C. *et al.* The haemangioblast generates haematopoietic cells through a haemogenic endothelium stage. *Nature* **457,** 892–895 (2009).
12. Nishikawa, S. *et al.* In vitro generation of lymphohematopoietic cells from endothelial cells purified from murine embryos. *Immunity* **8,** 761–769 (1998).
13. Boisset, J.-C. *et al.* In vivo imaging of haematopoietic cells emerging from the mouse aortic endothelium. *Nature* **464,** 116–120 (2010).
14. Bertrand, J. Y. *et al.* Haematopoietic stem cells derive directly from aortic endothelium during development. *Nature* **464,** 108–111 (2010).
15. Kissa, K. & Herbomel, P. Blood stem cells emerge from aortic endothelium by a novel type of cell transition. *Nature* **464,** 112–115 (2010).
16. Lam, E. Y. N., Hall, C. J., Crosier, P. S., Crosier, K. E. & Flores, M. V. Live imaging of Runx1 expression in the dorsal aorta tracks the emergence of blood progenitors from endothelial cells. *Blood* **116,** 909–914 (2010).
17. de Bruijn, M. F. T. R., Speck, N. A., Peeters, M. C. & Dzierzak, E. A. Definitive hematopoietic stem cells first develop within the major arterial regions of the mouse embryo. *EMBO J* **19,** 2465–2474 (2000).
18. Taoudi, S. & Medvinsky, A. J. Functional identification of the hematopoietic

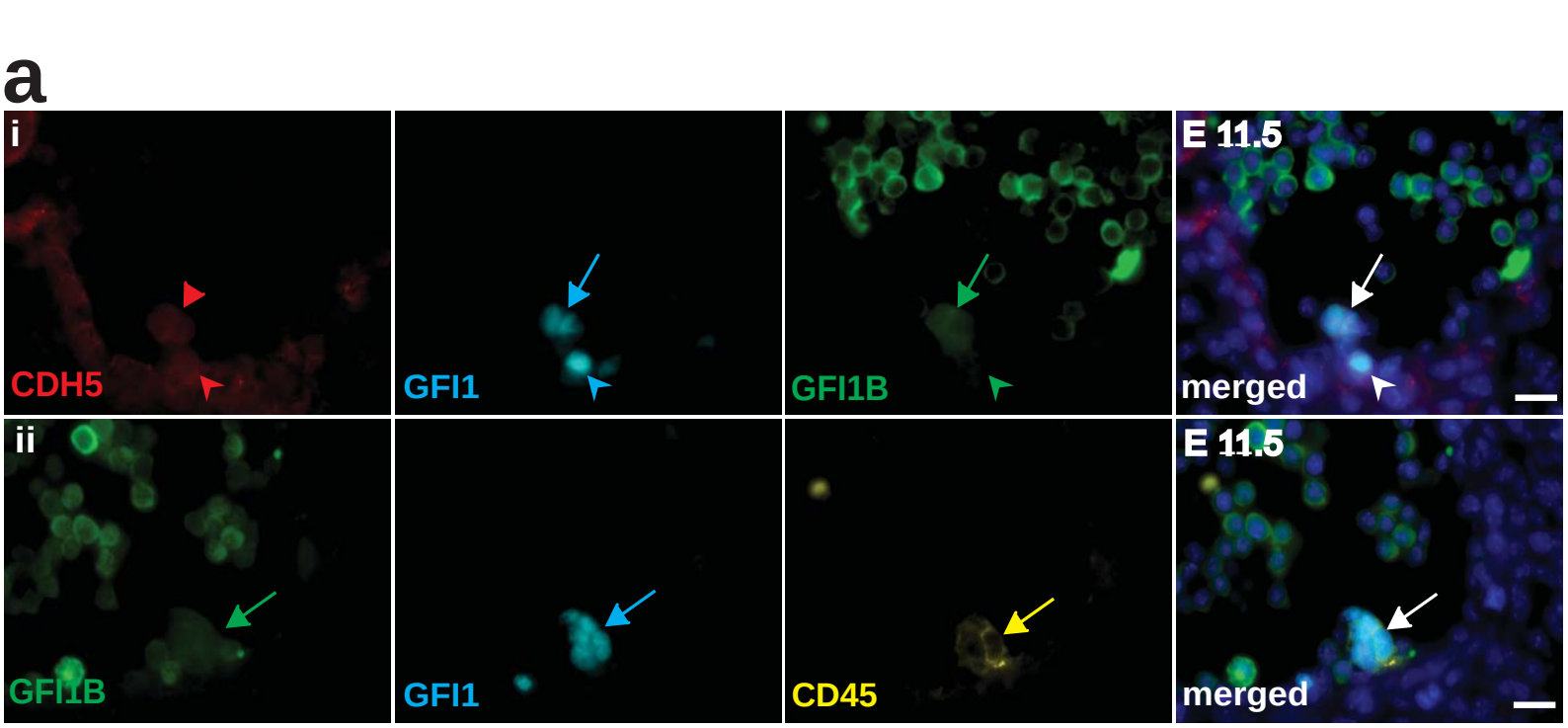
stem cell niche in the ventral domain of the embryonic dorsal aorta. *Proc Natl Acad Sci USA* **104,** 9399–9403 (2007).
19. Boisset, J.-C. *et al.* Progressive maturation toward hematopoietic stem cells in the mouse embryo aorta. *Blood* **125,** 465–469 (2015).
20. North, T. *et al.* Cbfa2 is required for the formation of intra-aortic hematopoietic clusters. *Development* **126,** 2563–2575 (1999).
21. Sroczynska, P., Lancrin, C., Kouskoff, V. & Lacaud, G. The differential activities of Runx1 promoters define milestones during embryonic hematopoiesis. *Blood* **114,** 5279–5289 (2009).
22. Lacaud, G. *et al.* Runx1 is essential for hematopoietic commitment at the hemangioblast stage of development in vitro. *Blood* **100,** 458–466 (2002).
23. North, T. E. *et al.* Runx1 expression marks long-term repopulating hematopoietic stem cells in the midgestation mouse embryo. *Immunity* **16,** 661–672 (2002).
24. North, T. E., Stacy, T., Matheny, C. J., Speck, N. A. & de Bruijn, M. F. T. R. Runx1 is expressed in adult mouse hematopoietic stem cells and differentiating myeloid and lymphoid cells, but not in maturing erythroid cells. *Stem Cells* **22,** 158–168 (2004).
25. Okuda, T., Hiebert, S. W. & Downing, J. R. AML1, the target of multiple chromosomal translocations in human leukemia, is essential for normal fetal liver hematopoiesis. *Cell* **84,** 321–330 (1996).
26. Wang, Q. *et al.* Disruption of the Cbfa2 gene causes necrosis and hemorrhaging in the central nervous system and blocks definitive hematopoiesis. *Proc Natl Acad Sci USA* **93,** 3444–3449 (1996).
27. Chen, M. J., Yokomizo, T., Zeigler, B. M., Dzierzak, E. A. & Speck, N. A. Runx1 is required for the endothelial to haematopoietic cell transition but not thereafter. *Nature* **457,** 887–891 (2009).
28. Lancrin, C. *et al.* GFI1 and GFI1B control the loss of endothelial identity of hemogenic endothelium during hematopoietic commitment. *Blood* **120,** 314–322 (2012).
29. Vassen, L., Okayama, T. & Möröy, T. Gfi1b: green fluorescent protein knock-in mice reveal a dynamic expression pattern of Gfi1b during hematopoiesis that is largely complementary to Gfi1. *Blood* **109,** 2356 (2007).
30. Rybtsov, S. *et al.* Hierarchical organization and early hematopoietic specification of the developing HSC lineage in the AGM region. *Journal of Experimental Medicine* **208,** 1305–1315 (2011).
31. Taoudi, S. *et al.* Extensive hematopoietic stem cell generation in the AGM region via maturation of VE-cadherin+CD45+ pre-definitive HSCs. *Cell Stem Cell* **3,** 99–108 (2008).
32. Yokomizo, T. & Dzierzak, E. A. Three-dimensional cartography of hematopoietic clusters in the vasculature of whole mouse embryos. *Development* **137,** 3651–3661 (2010).
33. Fiolka, K. *et al.* Gfi1 and Gfi1b act equivalently in haematopoiesis, but have distinct, non-overlapping functions in inner ear development. *EMBO Rep* **7,** 326–333 (2006).
34. Saleque, S., Orkin, S., Kim, J. & Rooke, H. M. Epigenetic regulation of hematopoietic differentiation by Gfi-1 and Gfi-1b is mediated by the cofactors CoREST and LSD1. *Molecular Cell* **27,** 562–572 (2007).
35. Harris, W. J. *et al.* The histone demethylase KDM1A sustains the oncogenic

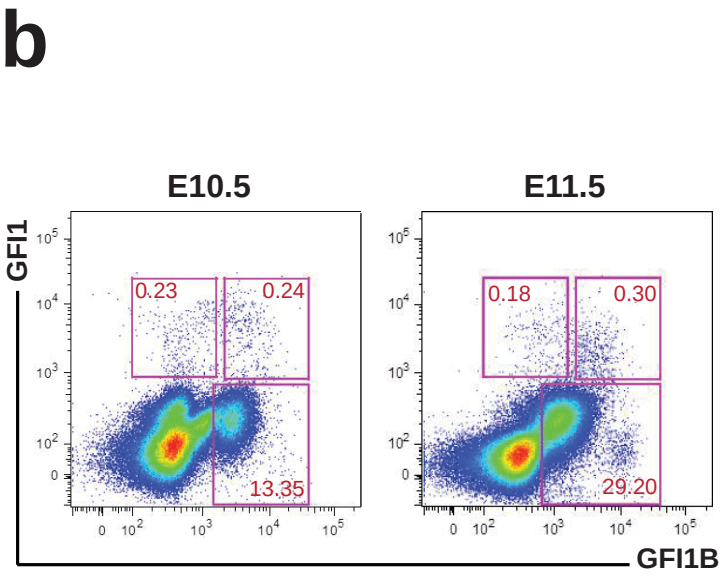
potential of MLL-AF9 leukemia stem cells. *Cancer Cell* **21,** 473–487 (2012).
36. Lancrin, C. *et al.* Blood cell generation from the hemangioblast. *J Mol Med* **88,** 167–172 (2010).
37. Foster, C. T. *et al.* Lysine-specific demethylase 1 regulates the embryonic transcriptome and CoREST stability. *Mol Cell Biol* **30,** 4851–4863 (2010).
38. Wang, J. *et al.* The lysine demethylase LSD1 (KDM1) is required for maintenance of global DNA methylation. *Nat Genet* **41,** 125–129 (2009).
39. Whyte, W. A. *et al.* Enhancer decommissioning by LSD1 during embryonic stem cell differentiation. *Nature* 1–5 (2012). doi:10.1038/nature10805
40. Vogel, M. J., Peric-Hupkes, D. & van Steensel, B. Detection of in vivo protein-DNA interactions using DamID in mammalian cells. *Nature Protocols* **2,** 1467–1478 (2007).
41. Lie-A-Ling, M. *et al.* RUNX1 positively regulates a cell adhesion and migration program in murine hemogenic endothelium prior to blood emergence. *Blood* (2014). doi:10.1182/blood-2014-04-572958
42. Lichtinger, M. *et al.* RUNX1 reshapes the epigenetic landscape at the onset of haematopoiesis. *EMBO J* 1–16 (2012). doi:10.1038/emboj.2012.275
43. Wang, L. & Dudek, S. M. Regulation of vascular permeability by sphingosine 1-phosphate. *Microvascular Research* **77,** 39–45 (2009).
44. Chen, M. J. *et al.* Erythroid/Myeloid Progenitors and Hematopoietic Stem Cells Originate from Distinct Populations of Endothelial Cells. *Stem Cell* **9,** 541–552 (2011).
45. Hadland, B. K. *et al.* A requirement for Notch1 distinguishes 2 phases of definitive hematopoiesis during development. *Blood* **104,** 3097–3105 (2004).
46. Kumano, K. *et al.* Notch1 but not Notch2 is essential for generating hematopoietic stem cells from endothelial cells. *Immunity* **18,** 699–711 (2003).
47. Bigas, A. & Robert-Moreno, L. The Notch pathway in the developing hematopoietic system. *Int. J. Dev. Biol* **54,** 1175–1188 (2010).
48. Guiu, J. *et al.* Hes repressors are essential regulators of hematopoietic stem cell development downstream of Notch signaling. *Journal of Experimental Medicine* **210,** 71–84 (2013).
49. Costa, G. *et al.* SOX7 regulates the expression of VE-cadherin in the haemogenic endothelium at the onset of haematopoietic development. *Development* **139,** 1587–1598 (2012).
50. Clarke, R. L. *et al.* The expression of Sox17 identifies and regulates haemogenic endothelium. *Nat Cell Biol* **15,** 1–10 (2013).
51. de Bruijn, M. F. T. R., Robin, C., Ottersbach, K., Sanchez, M.-J. & Dzierzak, E. A. Hematopoietic stem cells localize to the endothelial cell layer in the midgestation mouse aorta. *Immunity* **16,** 673–683 (2002).
52. Swiers, G. *et al.* Early dynamic fate changes in haemogenic endothelium characterized at the single-cell level. *Nature Communications* **4,** 2924 (2013).
53. Kim, W., Klarmann, K. D. & Keller, J. R. Gfi-1 regulates the erythroid transcription factor network through Id2 repression in murine hematopoietic progenitor cells. *Blood* **124,** 1586–1596 (2014).
54. Möröy, T. & Khandanpour, C. Growth factor independence 1 (Gfi1) as a regulator of lymphocyte development and activation. *Semin Immunol* **23,** 368–378 (2011).
55. Pearson, S., Cuvertino, S., Fleury, M., Lacaud, G. & Kouskoff, V. In Vivo Repopulating Activity Emerges at the Onset of Hematopoietic Specification

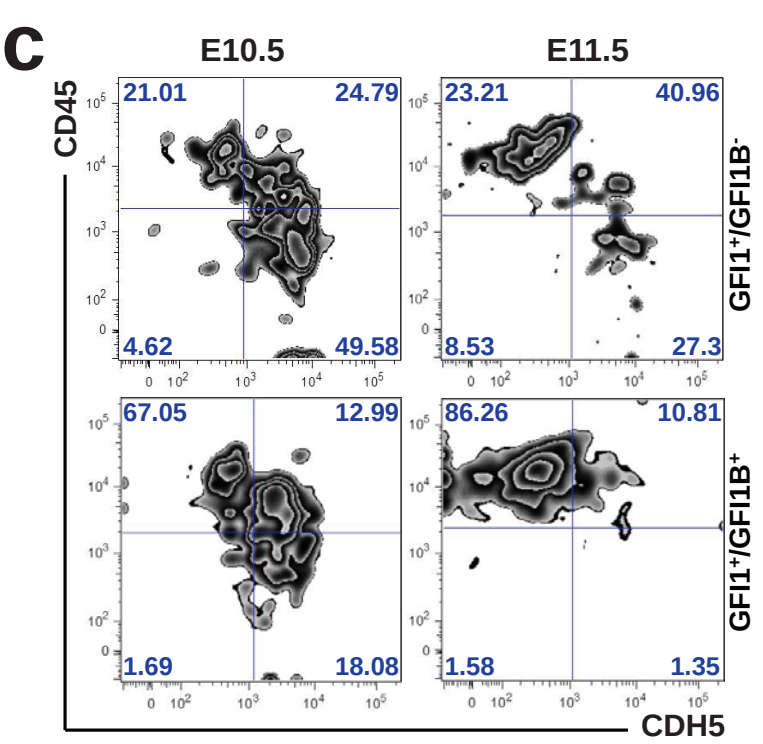
during Embryonic Stem Cell Differentiation. *Stem Cell Reports* **4,** 431–444 (2015).
56. Ledran, M. H. *et al.* Efficient hematopoietic differentiation of human embryonic stem cells on stromal cells derived from hematopoietic niches. *Cell Stem Cell* **3,** 85–98 (2008).
57. Wang, L. *et al.* Generation of hematopoietic repopulating cells from human embryonic stem cells independent of ectopic HOXB4 expression. *J Exp Med* **201,** 1603–1614 (2005).
58. Kyba, M., Perlingeiro, R. C. R. & Daley, G. Q. HoxB4 confers definitive lymphoid-myeloid engraftment potential on embryonic stem cell and yolk sac hematopoietic progenitors. *Cell* **109,** 29–37 (2002).
59. Pereira, C. F. *et al.* Induction of a hemogenic program in mouse fibroblasts. *Cell Stem Cell* **13,** 205–218 (2013).
60. Sandler, V. M. *et al.* Reprogramming human endothelial cells to haematopoietic cells requires vascular induction. *Nature* **511,** 312–318 (2014).
61. Batta, K., Florkowska, M., Kouskoff, V. & Lacaud, G. Direct reprogramming of murine fibroblasts to hematopoietic progenitor cells. *CellReports* **9,** 1871–1884 (2014).

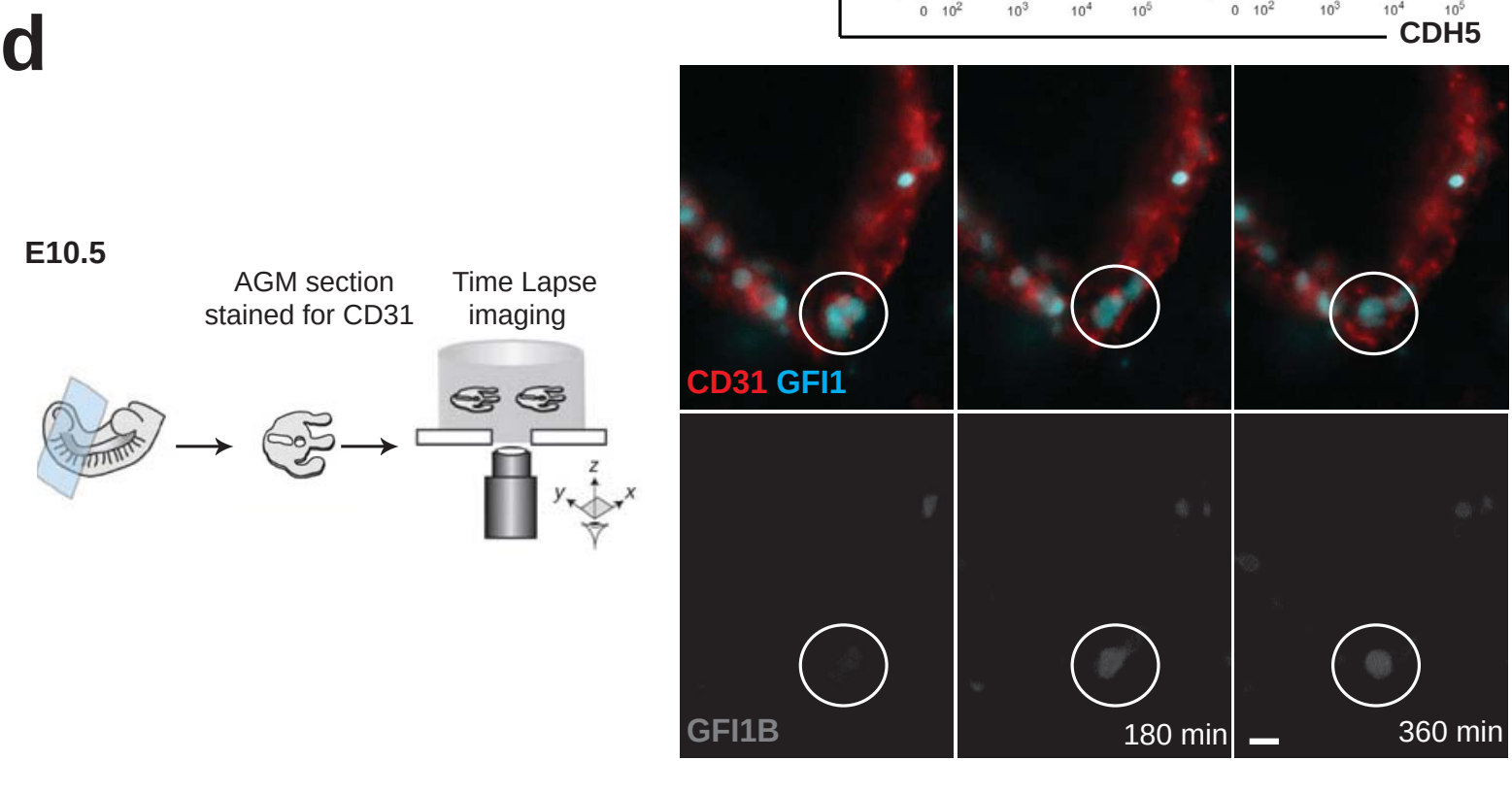

a
i
CD31
GFI1
GFI1B
E 9.5
merged
ii
CDH5
GFI1
GFI1B
E 10.5
merged
iii
CD31
GFI1
C-KIT
E 10.5
merged
iv
CDH5
GFI1
CD41
E 10.5
merged
v
GFI1B
GFI1
CD45
E10.5
merged

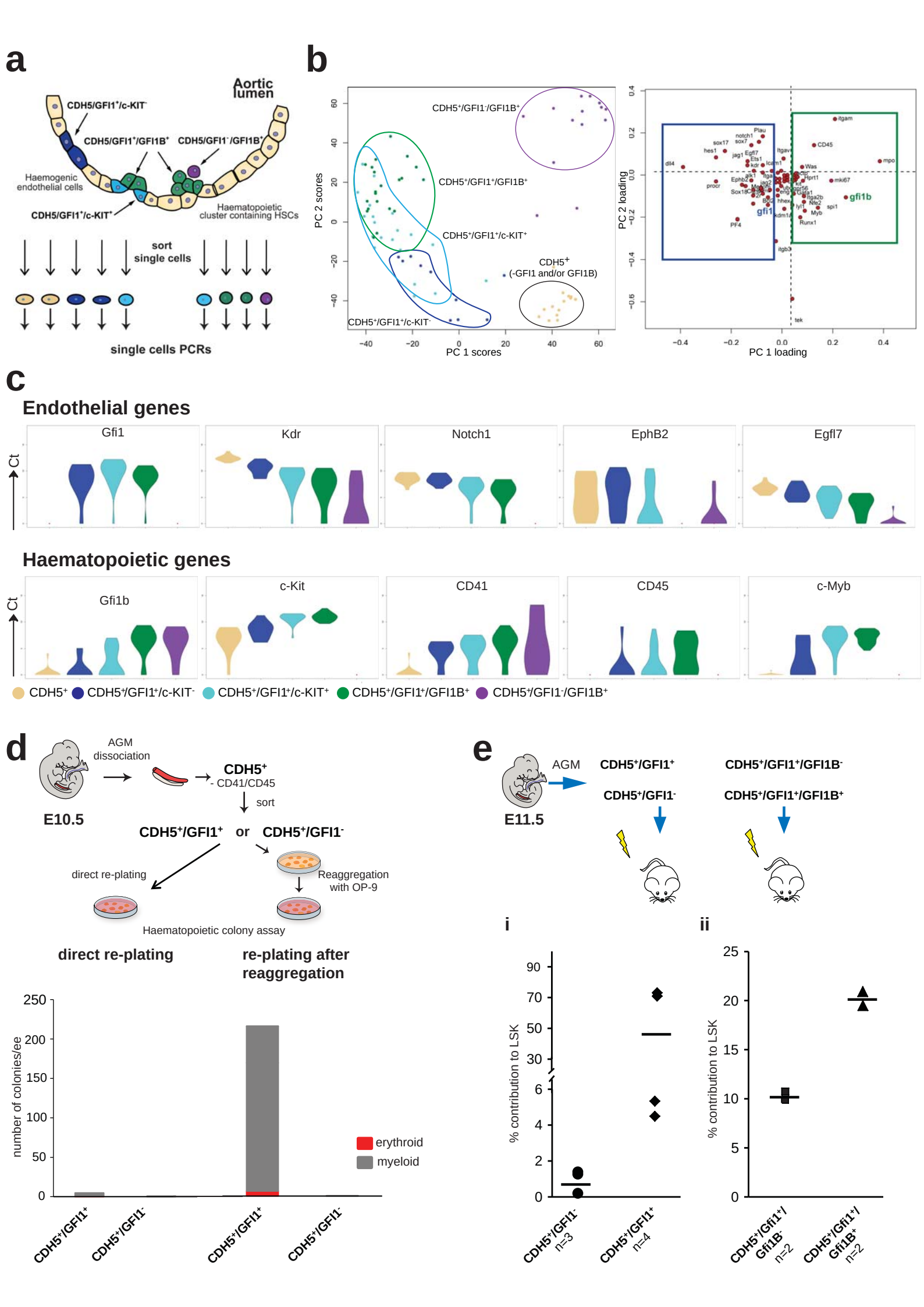

b
CD31
GFI1
RUNX1
CD31 RUNX1 DAPI
E10.5
CD31 GFI1 RUNX1 DAPI

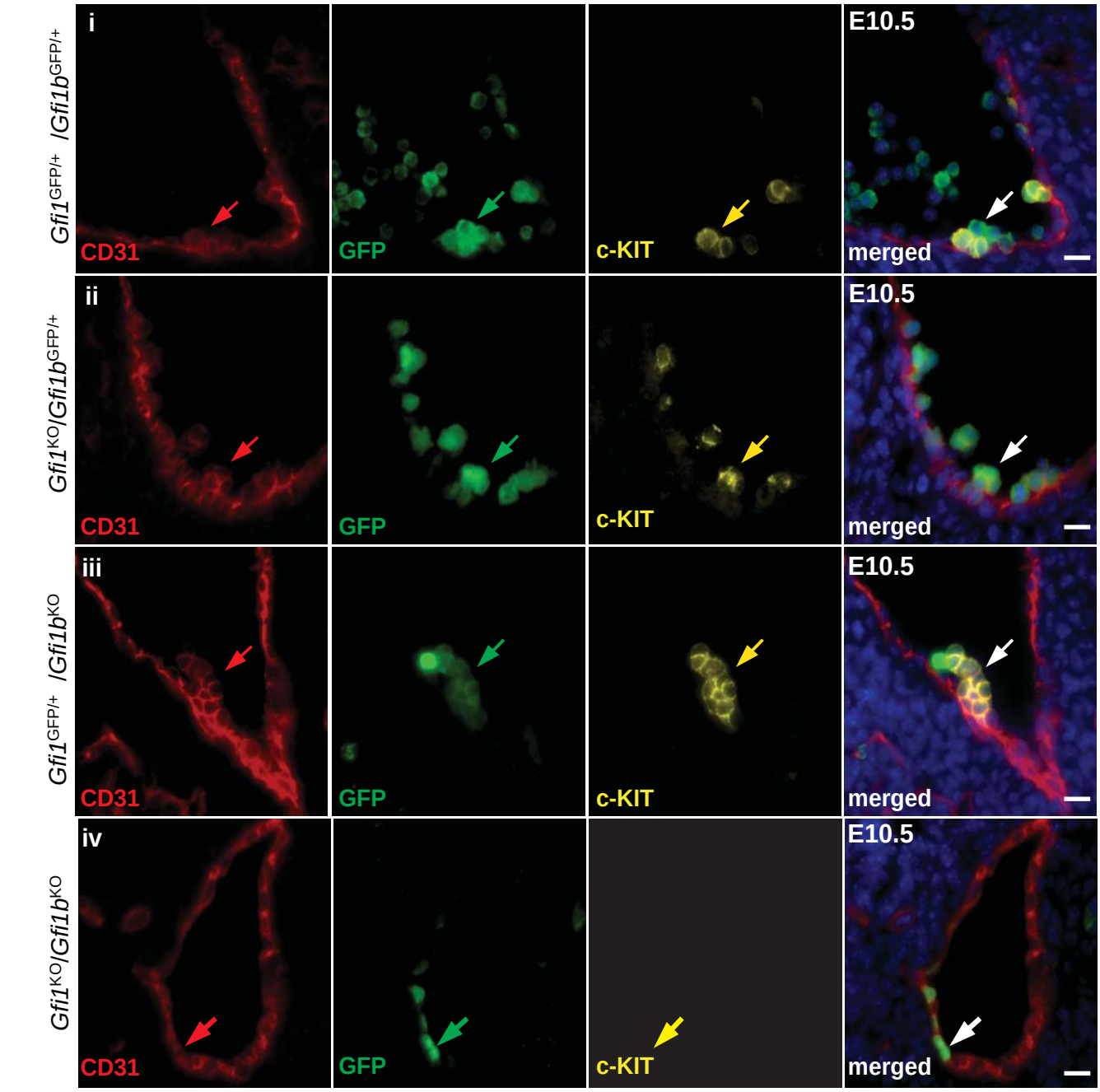

a
i
CDH5
GFI1
GFI1B
E 11.5
merged
ii
GFI1B
GFI1
CD45
E 11.5
merged

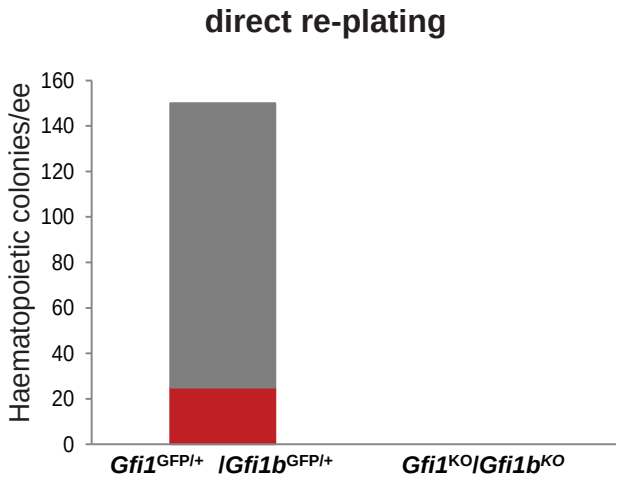

b
E10.5
E11.5
GFI1
GFI1B
0.23
0.24
13.35
0.18
0.30
29.20

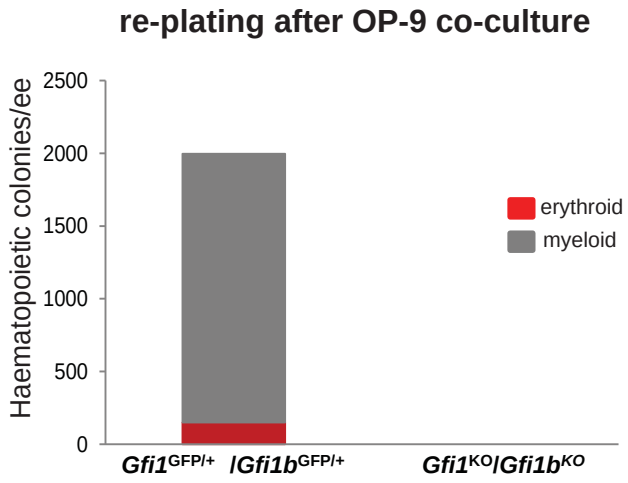

c
E10.5
E11.5
CD45
CDH5
GFI1+/GFI1B-
GFI1+/GFI1B+
21.01
24.79
4.62
49.58
23.21
40.96
8.53
27.3
67.05
12.99
1.69
18.08
86.26
10.81
1.58
1.35

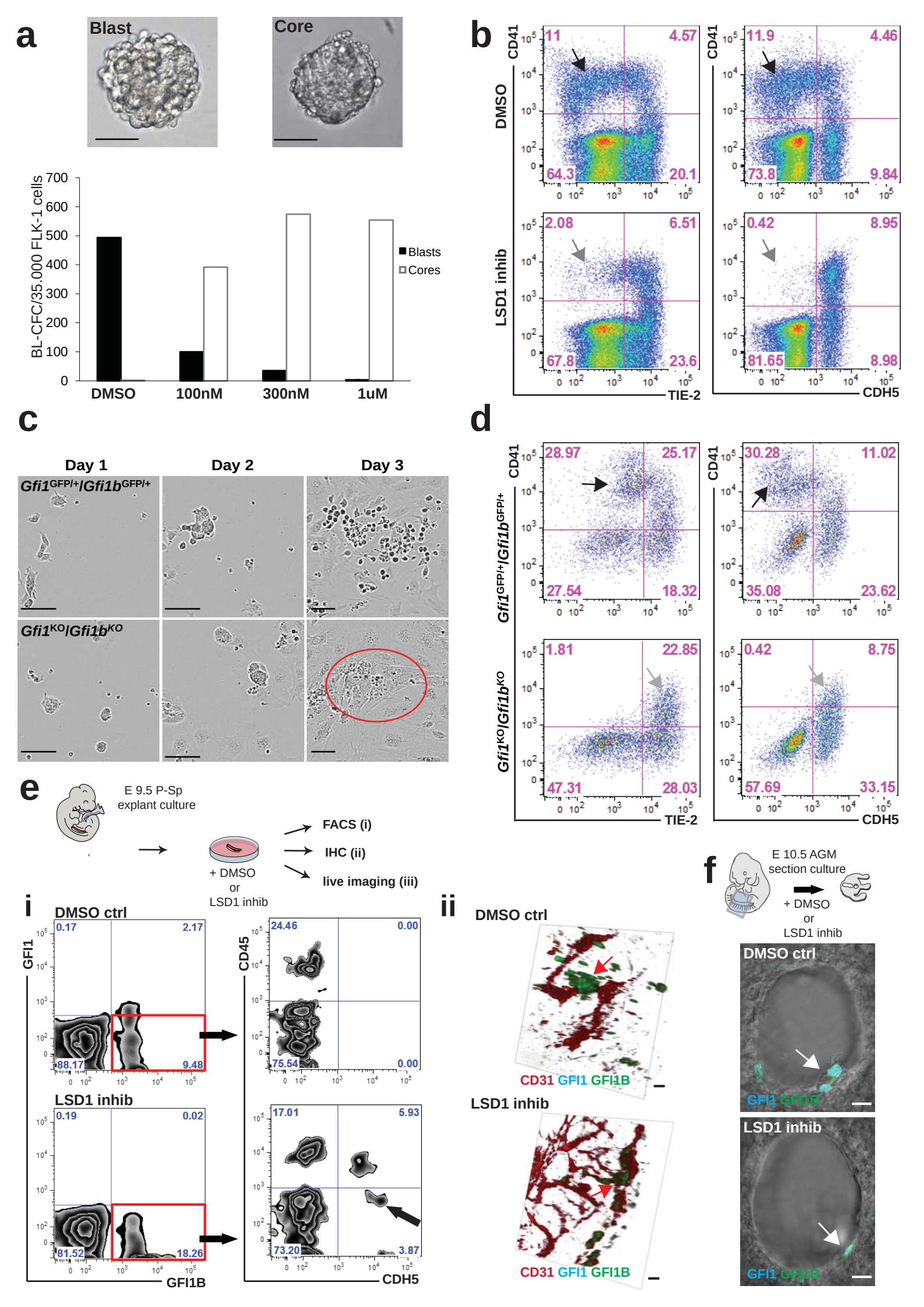

d
E10.5
AGM section
stained for CD31
Time Lapse
imaging
CD31 GFI1
GFI1B
180 min
360 min

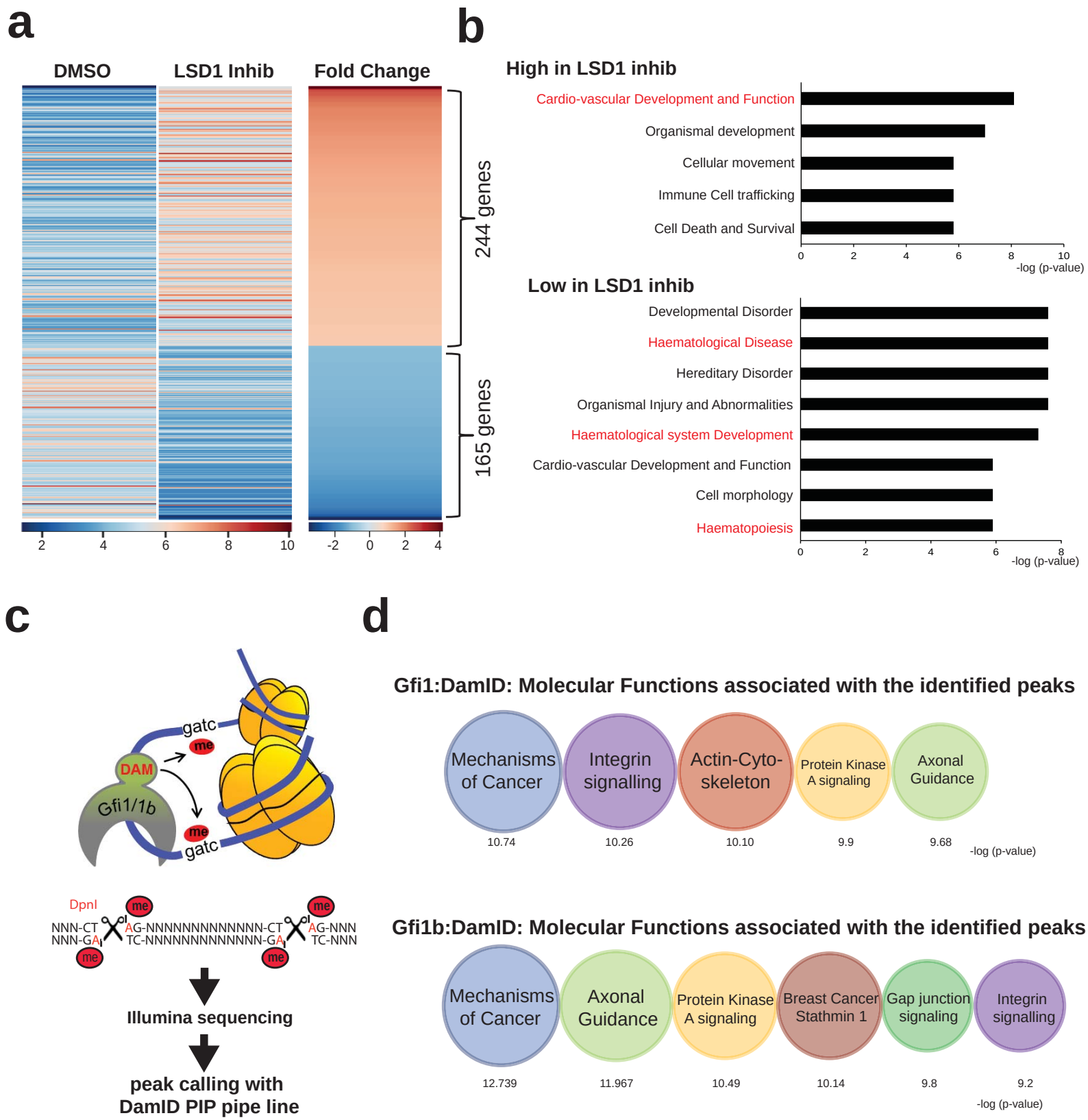

a
Aortic lumen
CDH5/GFI1+/c-KIT-
CDH5/GFI1+/GFI1B+
CDH5/GFI1-/GFI1B+
Haemogenic endothelial cells
CDH5/GFI1+/c-KIT+
Haematopoietic cluster containing HSCs
sort single cells
single cells PCRs
b
CDH5+/GFI1-/GFI1B+
CDH5+/GFI1+/GFI1B+
CDH5+/GFI1+/c-KIT+
CDH5+ (-GFI1 and/or GFI1B)
CDH5+/GFI1+/c-KIT-
PC 2 scores
PC 1 scores
PC 2 loading
PC 1 loading
gfi1
gfi1b
c
Endothelial genes
Gfi1
Kdr
Notch1
EphB2
Egfl7
Ct
Haematopoietic genes
Gfi1b
c-Kit
CD41
CD45
c-Myb
CDH5+
CDH5+/GFI1+/c-KIT-
CDH5+/GFI1+/c-KIT+
CDH5+/GFI1+/GFI1B+
CDH5+/GFI1-/GFI1B+
d
AGM dissociation
CDH5+ - CD41/CD45
sort
E10.5
CDH5+/GFI1+ or CDH5+/GFI1-
direct re-plating
Reaggregation with OP-9
Haematopoietic colony assay
direct re-plating
re-plating after reaggregation
number of colonies/ee
erythroid
myeloid
CDH5+/GFI1+
CDH5+/GFI1-
CDH5+/GFI1+
CDH5+/GFI1-
e
AGM
E11.5
CDH5+/GFI1+
CDH5+/GFI1-
CDH5+/GFI1+/GFI1B-
CDH5+/GFI1+/GFI1B+
i
ii
% contribution to LSK
CDH5+/GFI1- n=3
CDH5+/GFI1+ n=4
CDH5+/Gfi1+/ Gfi1B- n=2
CDH5+/Gfi1+/ Gfi1B+ n=2


a
i
ii
iii
iv
Gfi1GFP/+ /Gfi1bGFP/+
Gfi1KO/Gfi1bGFP/+
Gfi1GFP/+ /Gfi1bKO
Gfi1KO/Gfi1bKO
CD31
GFP
c-KIT
merged
E10.5

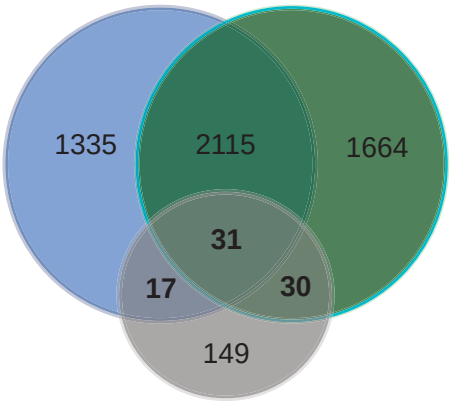

b
direct re-plating
Haematopoietic colonies/ee
0
20
40
60
80
100
120
140
160
Gfi1GFP/+
/Gfi1bGFP/+
Gfi1KO/Gfi1bKO


c
re-plating after OP-9 co-culture
Haematopoietic colonies/ee
0
500
1000
1500
2000
2500
erythroid
myeloid
Gfi1GFP/+
/Gfi1bGFP/+
Gfi1KO/Gfi1bKO


a
Blast
Core
BL-CFC/35.000 FLK-1 cells
700
600
500
400
300
200
100
0
DMSO
100nM
300nM
1uM
Blasts
Cores
b
CD41
DMSO
LSD1 inhib
11
4.57
64.3
20.1
11.9
4.46
73.8
9.84
2.08
6.51
67.8
23.6
0.42
8.95
81.65
8.98
TIE-2
CDH5
c
Day 1
Day 2
Day 3
Gfi1GFP/+/Gfi1bGFP/+
Gfi1KO/Gfi1bKO
d
CD41
Gfi1GFP/+/Gfi1bGFP/+
Gfi1KO/Gfi1bKO
28.97
25.17
27.54
18.32
30.28
11.02
35.08
23.62
1.81
22.85
47.31
28.03
0.42
8.75
57.69
33.15
TIE-2
CDH5
e
E 9.5 P-Sp
explant culture
+ DMSO
or
LSD1 inhib
FACS (i)
IHC (ii)
live imaging (iii)
i
DMSO ctrl
LSD1 inhib
GFI1
GFI1B
CD45
CDH5
0.17
2.17
88.17
9.48
24.46
0.00
75.54
0.00
0.19
0.02
81.52
18.26
17.01
5.93
73.20
3.87
ii
DMSO ctrl
LSD1 inhib
CD31 GFI1 GFI1B
f
E 10.5 AGM
section culture
+ DMSO
or
LSD1 inhib
DMSO ctrl
LSD1 inhib
GFI1 GFI1B


a
DMSO
LSD1 Inhib
Fold Change
244 genes
165 genes
2
4
6
8
10
-2
0
2
4
b
High in LSD1 inhib
Cardio-vascular Development and Function
Organismal development
Cellular movement
Immune Cell trafficking
Cell Death and Survival
0
2
4
6
8
10
-log (p-value)
Low in LSD1 inhib
Developmental Disorder
Haematological Disease
Hereditary Disorder
Organismal Injury and Abnormalities
Haematological system Development
Cardio-vascular Development and Function
Cell morphology
Haematopoiesis
0
2
4
6
8
-log (p-value)
c
gatc
me
DAM
Gfi1/1b
me
gatc
DpnI
me
me
NNN-CT
NNN-GA
AG-NNNNNNNNNNNNN-CT
TC-NNNNNNNNNNNNN-GA
AG-NNN
TC-NNN
me
me
Illumina sequencing
peak calling with
DamID PIP pipe line
d
Gfi1:DamID: Molecular Functions associated with the identified peaks
Mechanisms of Cancer
Integrin signalling
Actin-Cyto-skeleton
Protein Kinase A signaling
Axonal Guidance
10.74
10.26
10.10
9.9
9.68
-log (p-value)
Gfi1b:DamID: Molecular Functions associated with the identified peaks
Mechanisms of Cancer
Axonal Guidance
Protein Kinase A signaling
Breast Cancer Stathmin 1
Gap junction signaling
Integrin signalling
12.739
11.967
10.49
10.14
9.8
9.2
-log (p-value)

a



1700066B19Rik
ARHGEF15
CHST7
Ctla2a/Ctla2b
DUSP4
ENC1
GREB1L
HDX
HSD3B2
LGR5
LIN28A
NKAIN2
OIT3
PARP8
PRR16
TCF15
ZNF761

LSD1 inhib (high)

AIM2
AS3MT
ASAH2
ASB4
★CD93
CTTNBP2NL
CYP26A1
DPYSL3
★EGFL7
EPS8
★FBN1
★GATA2
HAND2
ID3
LPAR6
MCTS1
MDFIC
NOS3
NPL
PCDH12
PDYN
PTPRB
★RGL1
SCN11A
★SGPP1
SHE
★SOX7
STC1
★SYNE2
TBX3
WIPF3

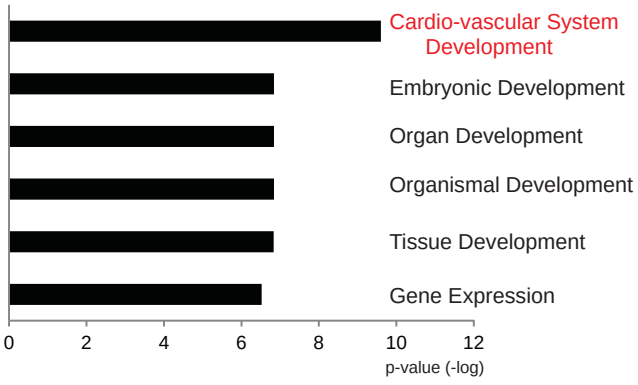
1600029D21Rik
2700046A07Rik
ACER2
ARAP3
BAALC
BC043934
BCL6B
C1orf54
C8orf4
CDH5
CYB5D1
★EDN1
ELTD1
EMX2
ESAM
FBN2
★GIMAP4
Gm266
GPM6A
HELT
HEY2
ICAM2
ID1
NPR1
POMC
RASGRP3
SALL1
SEC24D
TDGF1
THSD1

★ RBPJ binding motif in 100bp proximity of GFI1 and GFI1B peak

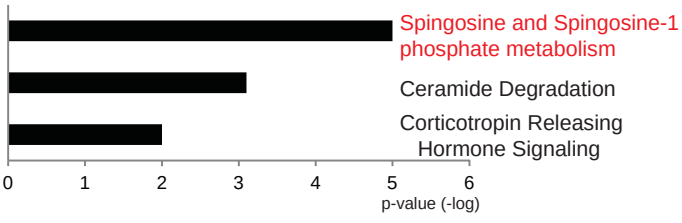
candidate(s) in RUNX1 ChIP-seq

b

Functional classification of the 78 candidates

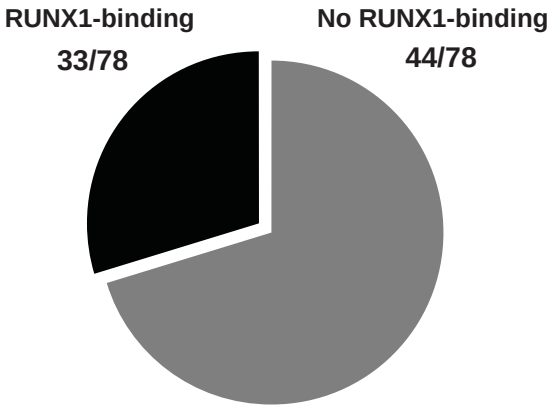


Enriched Pathways amongst the 78 candidates



c

Co-regulated by GFI1 and/or GFI1B and LSD1 in HE



d

Binding motifs 100bp up/down stream of GFI1:DAM and/or GFI1B:DAM peaks in the 78 final candidates:

RBPJ binding motif

p-value: 1e-10

e

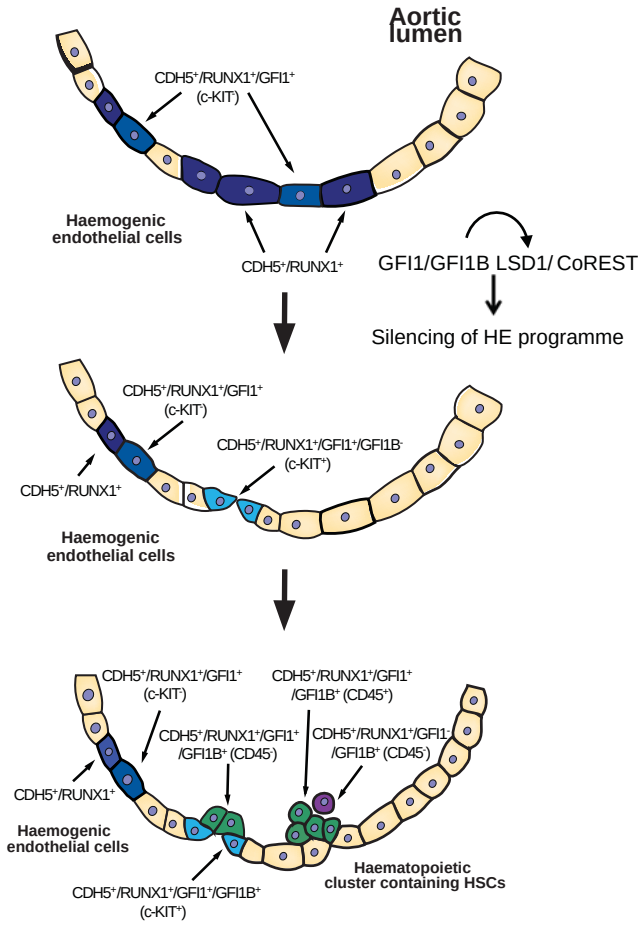


Figure legends for Supplementary Figures 1- 7

Supplementary Figure 1 *Gfi1*^{tomato} mouse line

(a) FACS analysis of *Gfi1*^{tomato} *and Gfi1*$^{GFP/+}$ bone marrow cells after staining for lymphoid (CD4/CD8) and myeloid (GR-1/MAC-1) markers. Bone marrow populations were gated into GFP^{-} and GFP^{+} for *Gfi1*$^{GFP/+}$ (top panel) and TOM^{-} and TOM^{+} for *Gfi1*^{tomato} (lower panel) fractions (representative FACS plot of one independent experiment). (**b**) IHC for CD31 (red), GFI1 (cyan) and GFI1B (green) on E10.5 AGM section counterstained with DAPI (representative images of 3 independent experiments). (**c**) Frequencies of GFI1^{+}/GFI1^{-} and GFI1^{+}/GFI1^{+} cells in the CDH5 compartment of E10.5 and E11.5 AGMs (AGMs from 11 independent litters, n= 6 for E10.5 and n=4 for E11.5 were analysed, p-values were determined by a two-tailed Student's t-test. Source data is available in Statistics source data file). (**d**) FACS analysis of GFI1 and GFI1B cells in the CDH5 compartment of E11.5 AGMs (representative FACS plot from one independent experiment). Scale bar = 10um.

Supplementary Figure 2 GFI1^{+}/GFI1B^{-} cells gain GFI1B expression *in vitro*

(**a**) CDH5^{+}/GFI1^{+}/GFI1B^{-} cells were FACS sorted from *Gfi1*^{tomato}*Gfi1*$^{GFP/+}$ E10.5 AGM and cultured on OP-9 stromal cells. Images at day 2 and 7 of *in vitro* cultures (representative images of n=4 independent experiments). (**b**) Day 7 cultures were analysed by FACS and indicated gates were used for sorting (**c**) The initial sorted E10.5 CDH5^{+}/GFI1^{+}/GFI1B^{-} AGM cells and the different cell populations isolated

after 7 days of their cultures (GFI1^{+}/GFI1B^{-}, GFI1^{+}/GFI1B^{+}, GFI1^{-}/GFI1B^{+}) were analysed for *Gfi1b* expression level by q-PCR (representative data from one experiment).

Supplementary Figure 3 Heatmap of all cells and genes analysed by single cell PCR

Heat map depicting the clustering all the single cells and genes analysed.

Supplementary Figure 4 E11.5 CDH5^{+}/GFI1^{+} cells contribution to recipients

(**a**) GFI1 and GFI1B expression mark all hematopoietic cells in the E10.5 AGM (**i**) schematic representation of the experimental design (**ii** and **iii**) E10.5 AGM cells were FACS sorted and re-plated into haematopoietic assay either before (one independent experiment with embryos from one litter) (**ii**) or after (one independent experiment with embryos from one litter) (**iii**) co-culture step with OP-9. Colonies were scored 9-11 days later. (**b**) FACS analysis of recipients transplanted with either CDH5^{+}/GFI1^{+} or CDH5^{+}/GFI1^{-} E11.5 AGM cells (CD45.1/CD45.2 double positives) at 17 weeks after transplantation. Bone marrow LSK compartments of recipients (CD45.1) were analysed for CD45.1 and CD45.2 (representative FACS plots of one independent experiment). (**c**) Donor contribution to haematopoietic lineages determined by sub-gating on CD45.1/CD45.2 double positive population in the recipient. Donor contribution to T cells (CD4/CD8) in the spleen and thymus, B cells (IgM/B220), myeloid cells (GR-1/MAC-1) and erythroid cells (CD71/TER119) in

different haematopoietic organs are shown (representative FACS plots of one independent experiment).

Supplementary Figure 5 GFI1 or GFI1B single knock out embryos can generate IAHC

(**a**) IHC on E10.5 *GFI1*^{KO}*GFI1B*$^{+/+}$ and *GFI1*$^{+/+}$*GFI1B*^{KO} AGM section for CD31, GFP and c-KIT (counterstained with DAPI) (representative images of one independent experiment with embryos form different litters). (**b**) *In situ* hybridisation for *Gfi1b* (red dots) and IHC for CD31 (brown) on *GFI1*$^{GFP/+}$*GFI1B*$^{GFP/+}$ and *GFI1*^{KO}*GFI1B*$^{GFP/+}$ E10.5 embryo sections. (**c**) qPCRs on E10.5 AGM cells. (**i**) qPCR for *Gfi1* and *Gfi1b* expression in AGM cell lysate of *GFI1*$^{GFP/+}$*GFI1B*$^{GFP/+}$ (het/het), *GFI1*$^{GFP/GFP}$*GFI1B*$^{GFP/+}$ (KO/het) and *GFI1*$^{GFP/+}$*GFI1B*$^{GFP/GFP}$ (het/KO) embryos (representative data from one independent experiment with embryos from the same litter). (**ii**) PCR for *Gfi1* and *Gfi1b* on sorted CDH5^{+}/GFP^{+}/c-KIT^{-} HE cells of *GFI1*$^{GFP/+}$*GFI1B*$^{GFP/+}$ (het/het) and *GFI1*$^{GFP/GFP}$*GFI1B*$^{GFP/+}$ (KO/het) embryos (one independent experiment with embryos from the same litter). Scale bar=10um

Supplementary Figure 6 Conditional *Lsd1* knock-out recapitulates LSD1 inhibition phenotype and leads to decrease in proliferation and apoptosis

(**a**) FLK1^{+} cells from the conditional *Lsd1*$^{\Delta/lox}$ line were isolated from day 3 EBs and cultured as monolayer. FACS analysis on 3 consecutive days with staining for TIE-2, CDH5 and CD41 are shown. *Lsd1*$^{\Delta/lox}$ FLK1+ cells were either cultured with control

ETOH or with 1uM of 4OHT (in ETOH) to induce the activity of *Cre-ERT2* and generate *Lsd1*$^{\Delta/\Delta}$ (representative FACS plots of 3 independent experiments) (**b**) EdU assay on Day 3 control (*Lsd1*$^{\Delta/lox}$) or *Lsd1* deleted (*Lsd1* $^{\Delta/\Delta}$) Blast cultures (Representative FACS plots of one independent experiment). (**c**) Annexin5/7-AAD staining on Day 3 control (*Lsd1*$^{\Delta/lox}$) or *Lsd1* deleted (*Lsd1*$^{\Delta/\Delta}$) Li-Blast cultures. Additionally, the cultures were stained with CD41 to differentiate between HE and non-HE cells.

Supplementary Figure 7 LSD1 is ubiquitously expressed and its inhibition leads to loss of EHT in E10.5 AGM.

(**a)** IHC for CD31, tomato and LSD1 on E10.5 AGM sections of *Gfi1*^{tomato} embryos (counterstained with DAPI). (**b**) IHC for CDH5, GFI1 and CD45 on E9.5 P-Sp explants. (**c**) Single channel control images from live imagining of E10.5 *Gfi1*^{tomato}*Gfi1*$^{GFP/+}$ AGM slices wither treated with DMSO (ctrl) or 500nM of the LSD1 inhibitor. (**di-iii**) Screen shots of the *Oit-3*, *Gata-2* and *Hey-2* locus from the UCSC browser showing examples of GFI1:DAM (Cyan) and GFI1B:DAM (green) binding enrichment after sequencing. Region of interest is marked with a black box and specific peaks detected are highlighted with black bars below.

Supplementary video 1

Time lapse of culture of sections of E10.5 *Gfi1*^{Tomato}*/Gfi1b*^{GFP} AGMs.

Supplementary video 2

Time-lapse imaging of FLK1^{+} cells blast monolayer cultures treated with DMSO (DMSO Ctrl).

Supplementary video 3

Time-lapse imaging of FLK1^{+} cells blast monolayer cultures treated with 300nM of the LSD1 inhibitor (LSD1 inhib).

Supplementary video 4

3D re-construction of z-stacks (4.2um) of E9.5 P-Sp explants (2 days cultures) treated with DMSO and stained for CD31 (red), GFI1 (cyan) and GFI1B (green).

Supplementary video 5

3D re-construction of z-stacks (4.1um) of E9.5 P-Sp explants (2 days cultures) treated with the LSD1 inhibitor (300nM) and stained for CD31 (red), GFI1 (cyan) and GFI1B (green).

Supplementary movie 6

Time-lapse imaging of FLK1^{+} cells from the c*Lsd1* ES line in Liquid Blast culture treated with EtOH (*Lsd1*$^{\Delta/lox}$).

Supplementary movie 7

Time-lapse imaging of FLK1^{+} cells from the c*Lsd1* ES line in Liquid Blast culture treated with 4-OHT (*Lsd1*$^{\Delta/\Delta}$).

Supplementary table 1

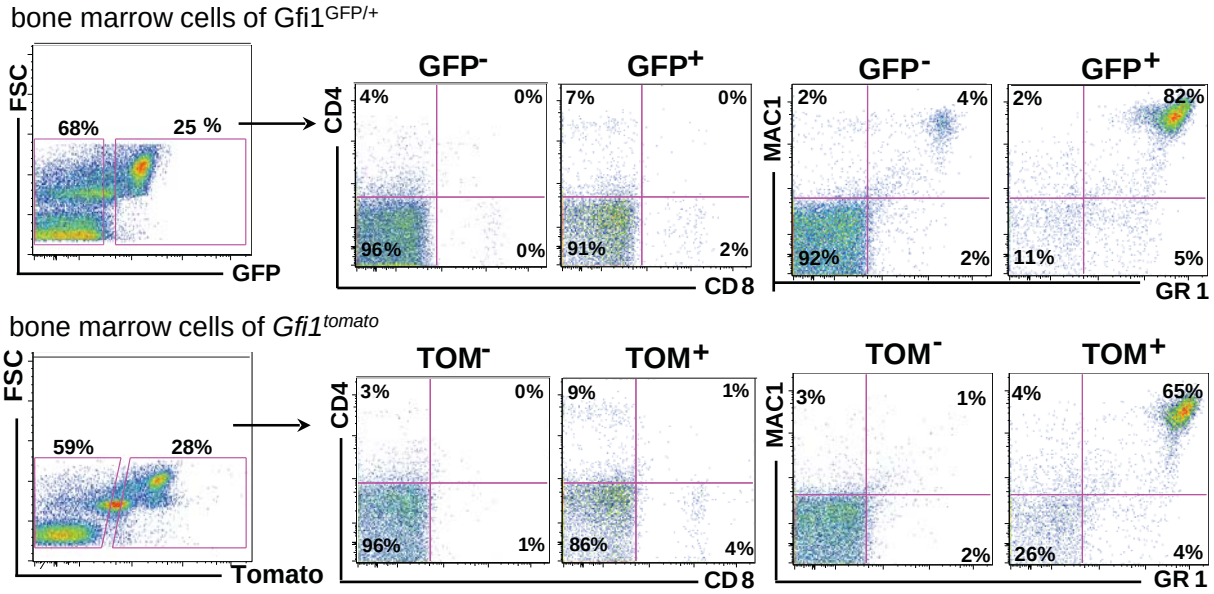
List of primary and secondary antibodies and their dilution used in this study

Supplementary table 2

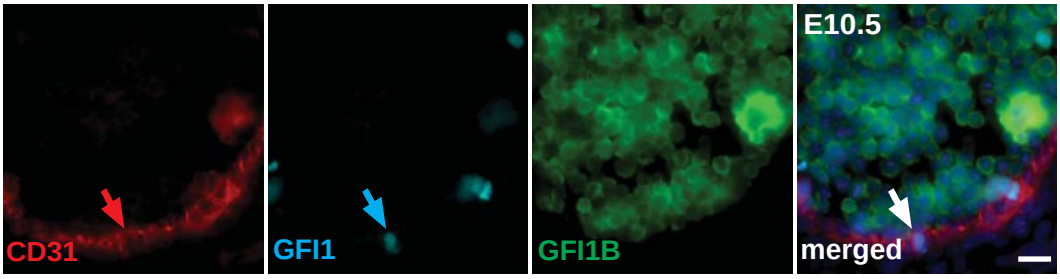
Raw data file from the single cell q-PCR Biomark platform showing the Ct values for each cell and gene analyzed

S1

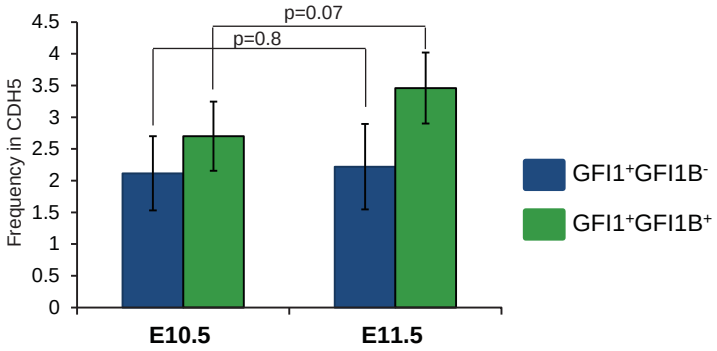
a


bone marrow cells of Gfi1GFP/+
FSC
68%
25 %
GFP
GFP-
GFP+
CD4
4%
0%
96%
0%
7%
0%
91%
2%
CD8
MAC1
2%
4%
92%
2%
2%
82%
11%
5%
GR 1
bone marrow cells of Gfi1tomato
FSC
59%
28%
Tomato
TOM-
TOM+
CD4
3%
0%
96%
1%
9%
1%
86%
4%
CD8
MAC1
3%
1%
2%
4%
65%
26%
4%
GR 1

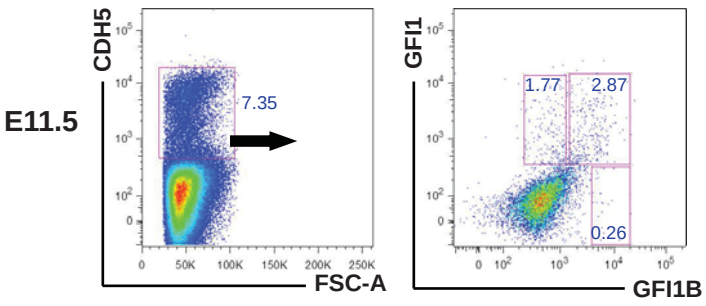
b


CD31
GFI1
GFI1B
E10.5
merged

c

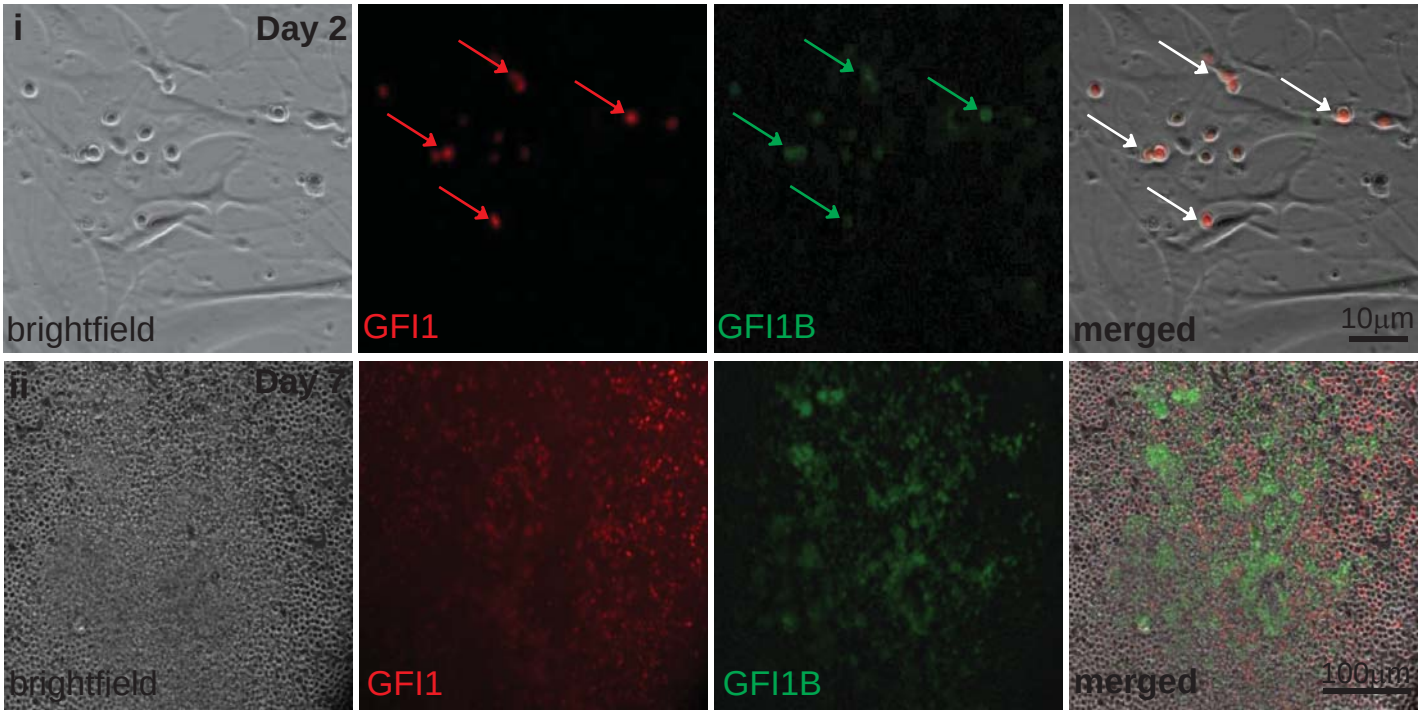

p=0.07
p=0.8
Frequency in CDH5
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
E10.5
E11.5
GFI1+GFI1B-
GFI1+GFI1B+

d


E11.5
CDH5
7.35
FSC-A
GFI1
1.77
2.87
0.26
GFI1B

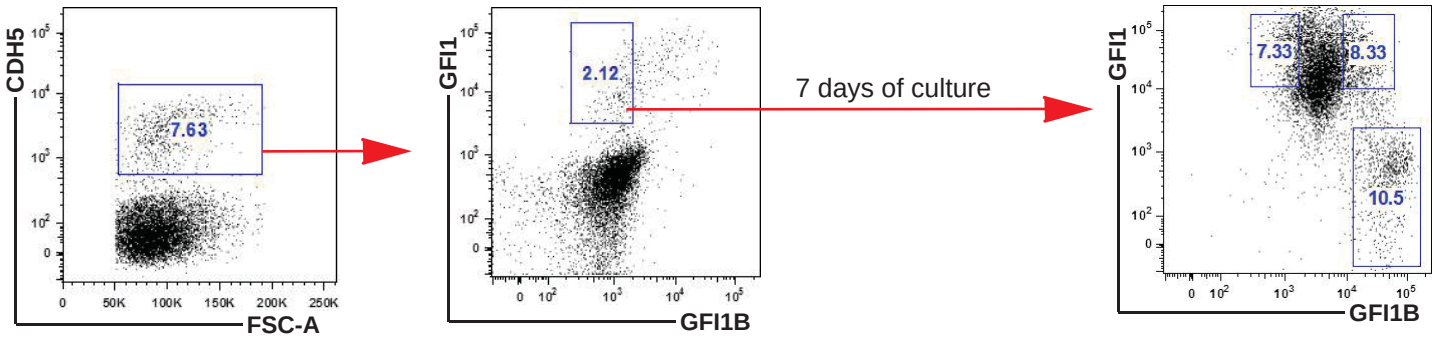
S2

a

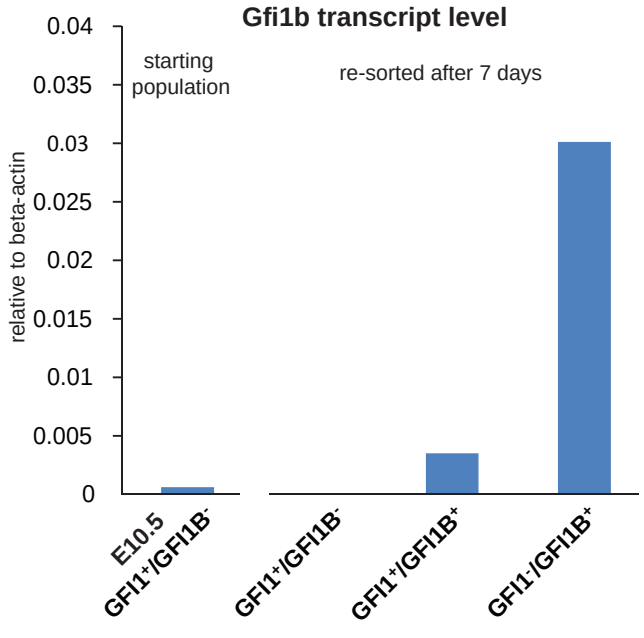


b

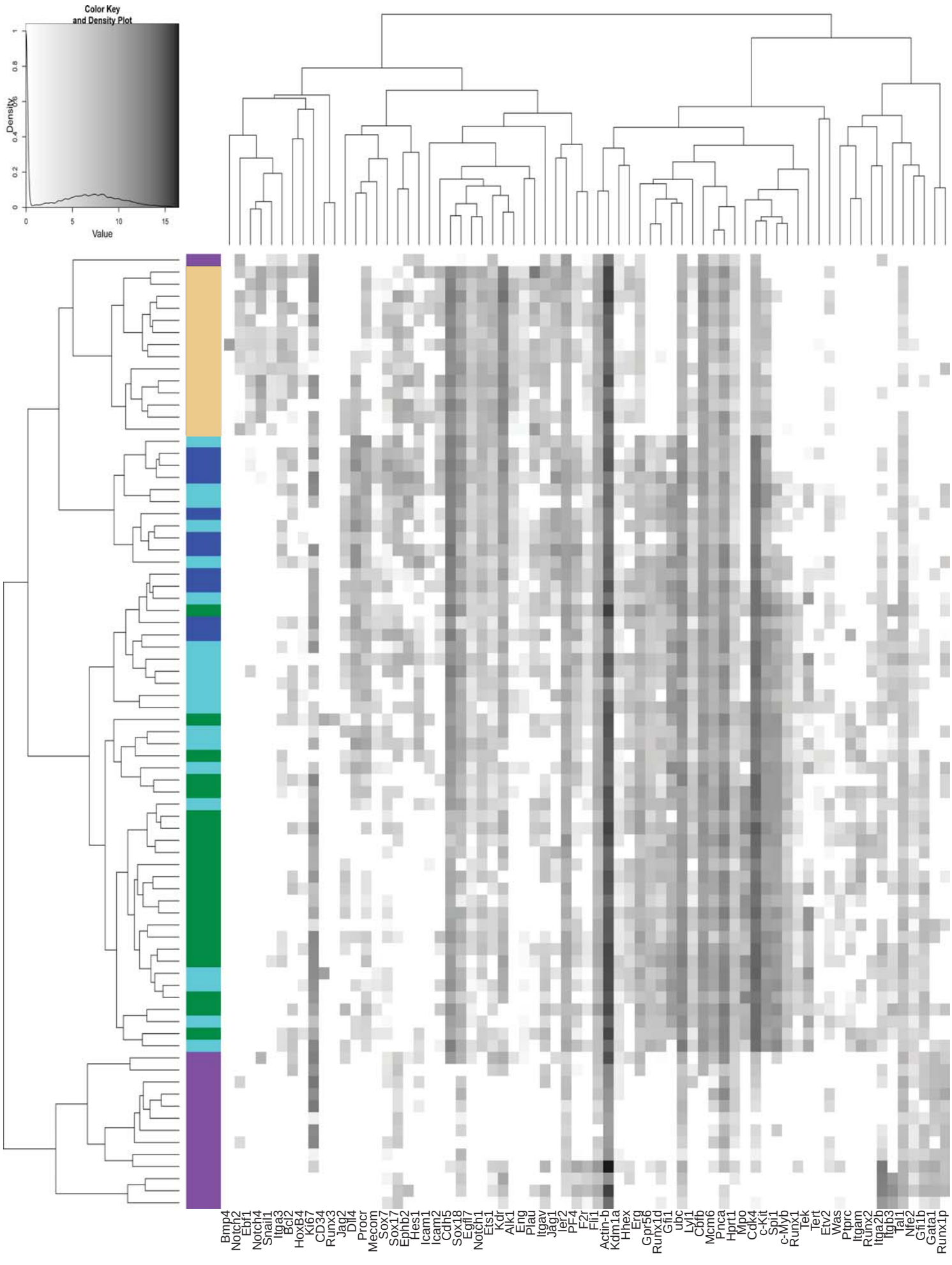
E10.5 *Gfi1*^{tomato} *Gfi1b*$^{+/GFP}$ AGMs



c



S3

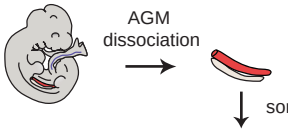

Color Key
and Density Plot
Density
Value
Bmp4
Notch2
Ebf1
Notch4
Snail1
Itga3
Bcl2
HoxB4
Ki67
CD34
Runx3
Jag2
Dll4
Procr
Mecom
Sox7
Sox17
Ephb2
Hes1
Icam1
Icam2
Cdh5
Sox18
Egfl7
Notch1
Ets1
Kdr
Alk1
Eng
Plau
Itgav
Jag1
Ier2
PF4
F2r
Fli1
Actin-b
Kdm1a
Hhex
Erg
Gpr56
Runx1d
Gfi1
ubc
Lyl1
Cbfb
Mcm6
Pnca
Hprt1
Mpo
Cdk4
c-Kit
Spi1
c-Myb
Runx1
Tek
Tert
Etv2
Was
Ptprc
Itgam
Runx2
Itga2b
Itgb3
Tal1
Nfe2
Gfi1b
Gata1
Runx1p
CDH5+
CDH5+/GFI1+/c-KIT-
CDH5+/GFI1+/c-KIT+
CDH5+/GFI1+/GFI1B+
CDH5+/GFI1-/GFI1B+

S4

a

E10.5

i



GFI1$^-$/GFI1B$^-$ GFI1$^+$/GFI1B$^-$ GFI1$^+$/GFI1B$^+$ GFI1$^-$/GFI1B$^+$

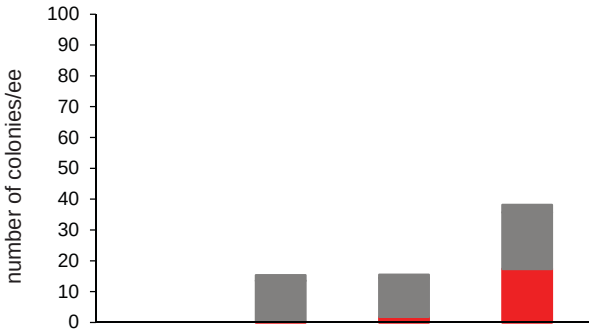
direct re-plating or after OP-9 Co-culture

Haematopoietic colony assay

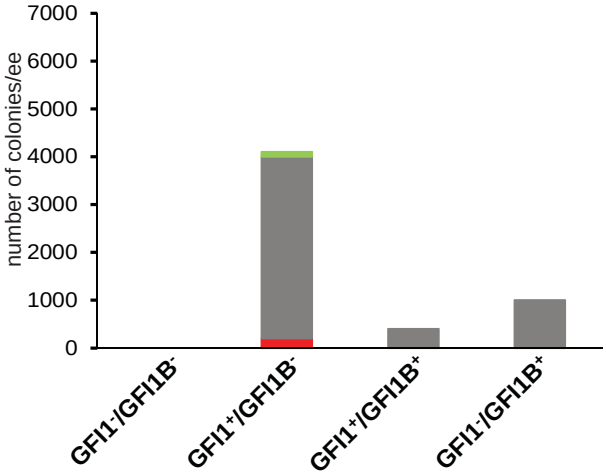
direct re-plating

re-plating after co-culture with OP-9

ii



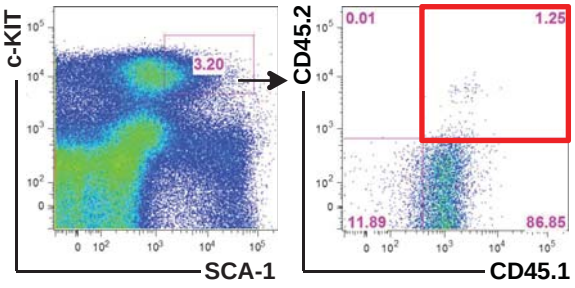
iii



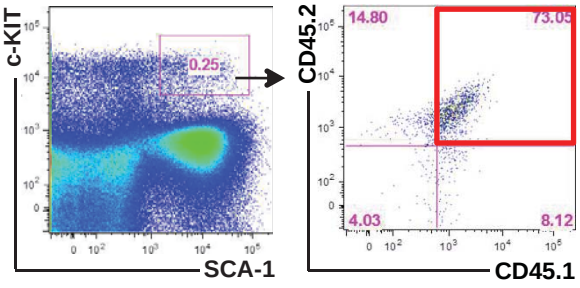
b

Donor contribution to the LSK compartment in recipient at 17 weeks post transplantation

CDH5$^+$/GFI1$^-$



CDH5$^+$/GFI1$^+$

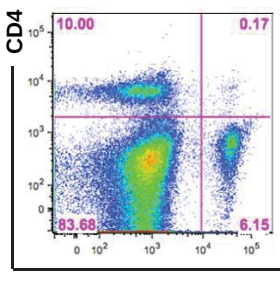


c

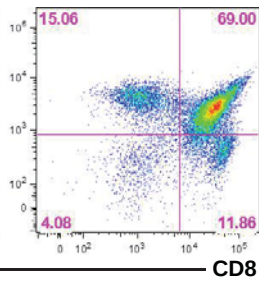
Multi- lineage contribution: gated on CD45.1$^+$/CD45.2$^+$ in recipient

Lymphoid lineage

Spleen

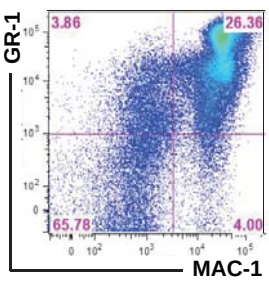


Thymus

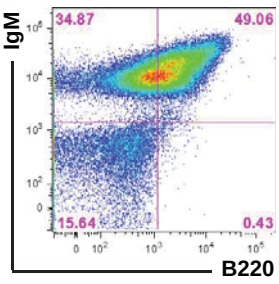


Myeloid lineage

bone marrow

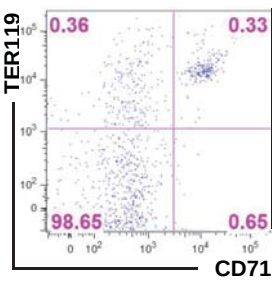


Spleen

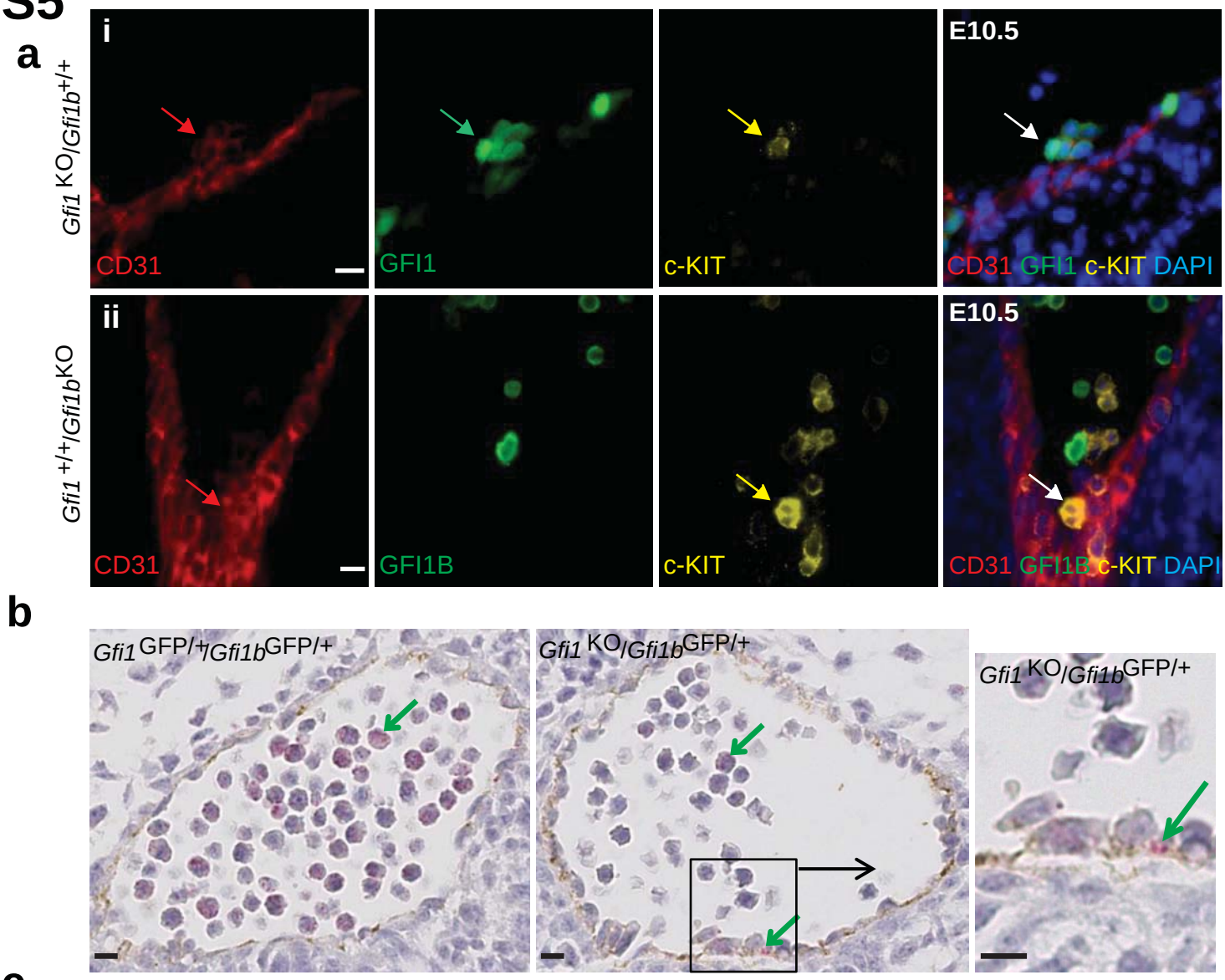


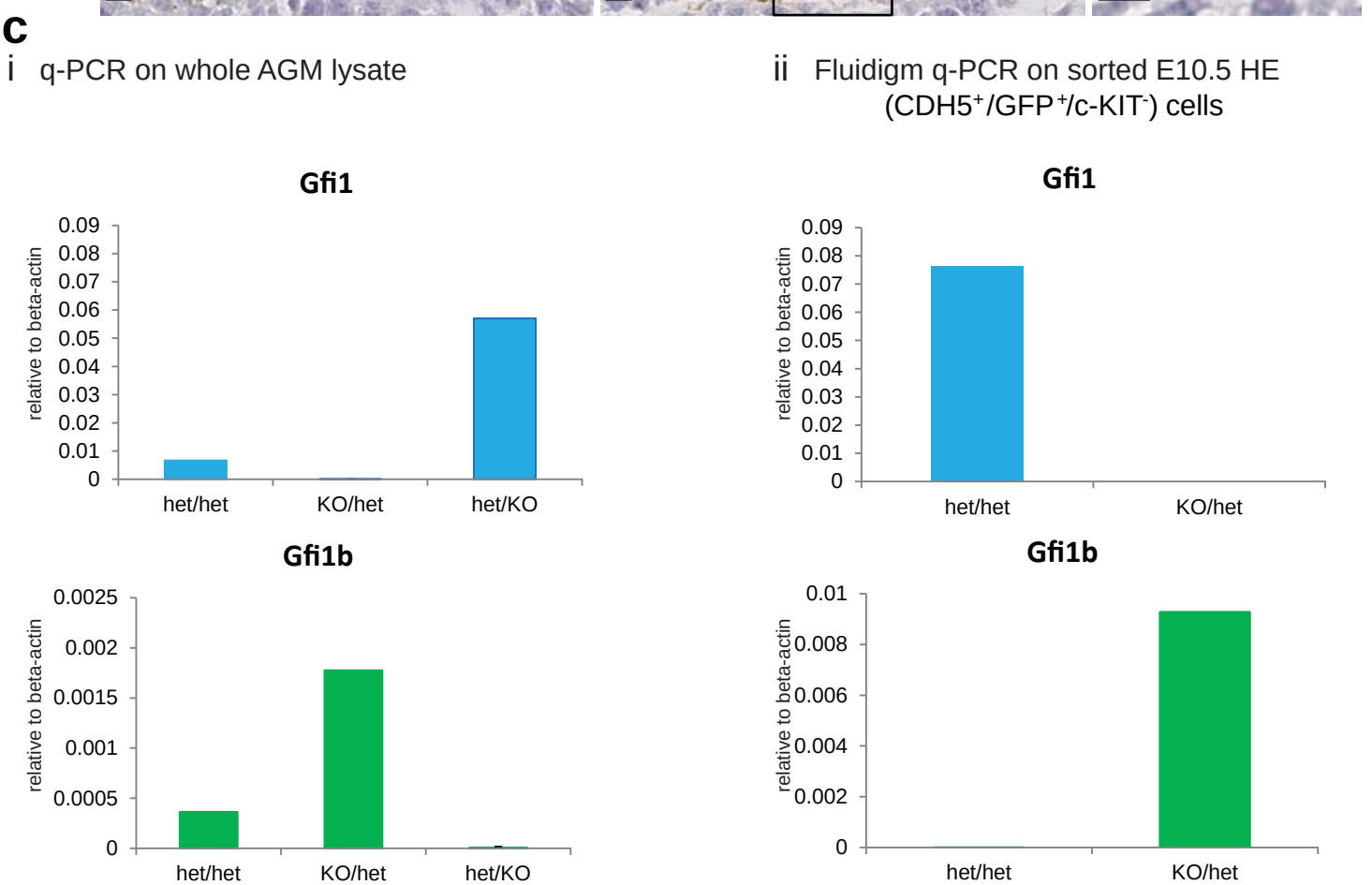
Erythroid lineage

bone marrow



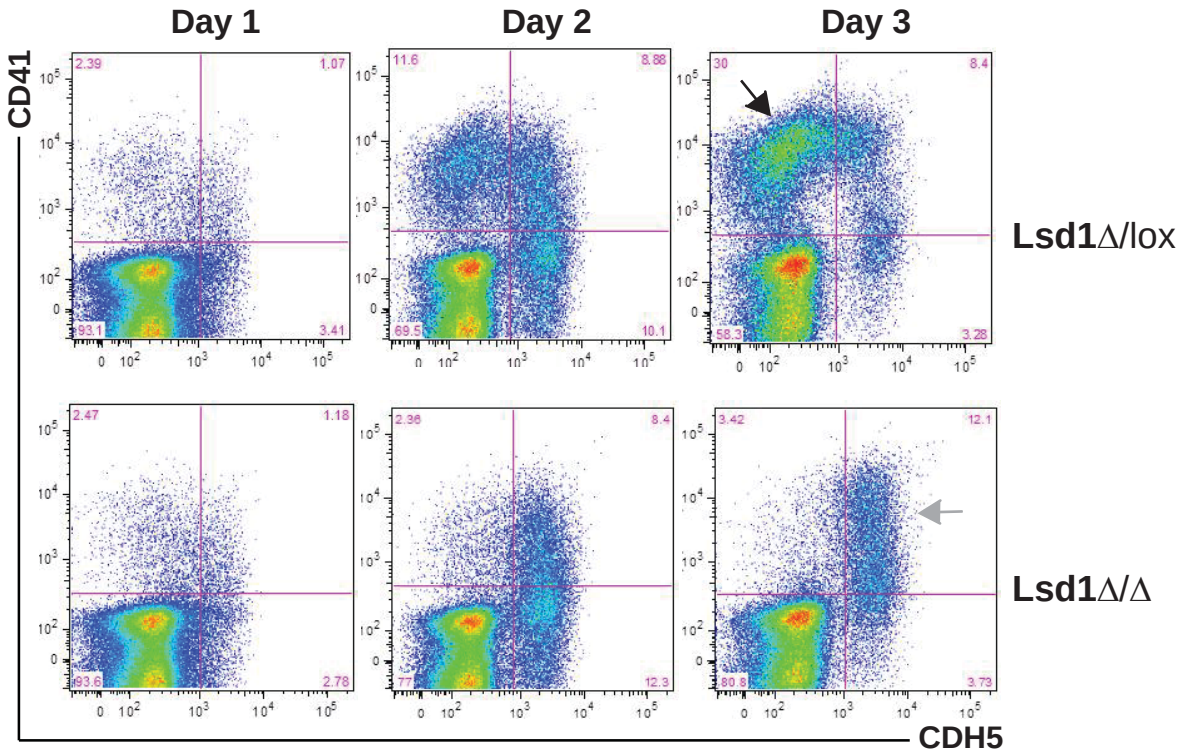
S5


a
i
Gfi1 KO/Gfi1b+/+
CD31
GFI1
c-KIT
E10.5
CD31 GFI1 c-KIT DAPI
ii
Gfi1 +/+/Gfi1bKO
CD31
GFI1B
c-KIT
E10.5
CD31 GFI1B c-KIT DAPI
b
Gfi1 GFP/+/Gfi1b GFP/+
Gfi1 KO/Gfi1b GFP/+
Gfi1 KO/Gfi1b GFP/+

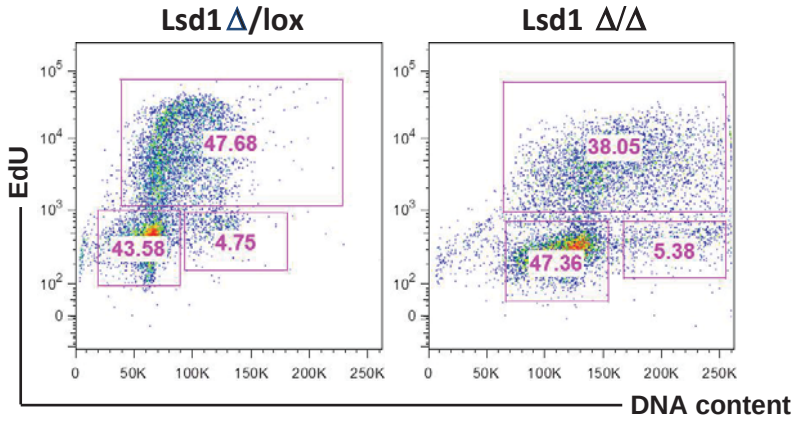

c
i q-PCR on whole AGM lysate
ii Fluidigm q-PCR on sorted E10.5 HE (CDH5+/GFP+/c-KIT-) cells
Gfi1
relative to beta-actin
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
het/het
KO/het
het/KO
Gfi1
relative to beta-actin
het/het
KO/het
Gfi1b
relative to beta-actin
0.0025
0.002
0.0015
0.001
0.0005
0
het/het
KO/het
het/KO
Gfi1b
relative to beta-actin
0.01
0.008
0.006
0.004
0.002
0
het/het
KO/het

S6

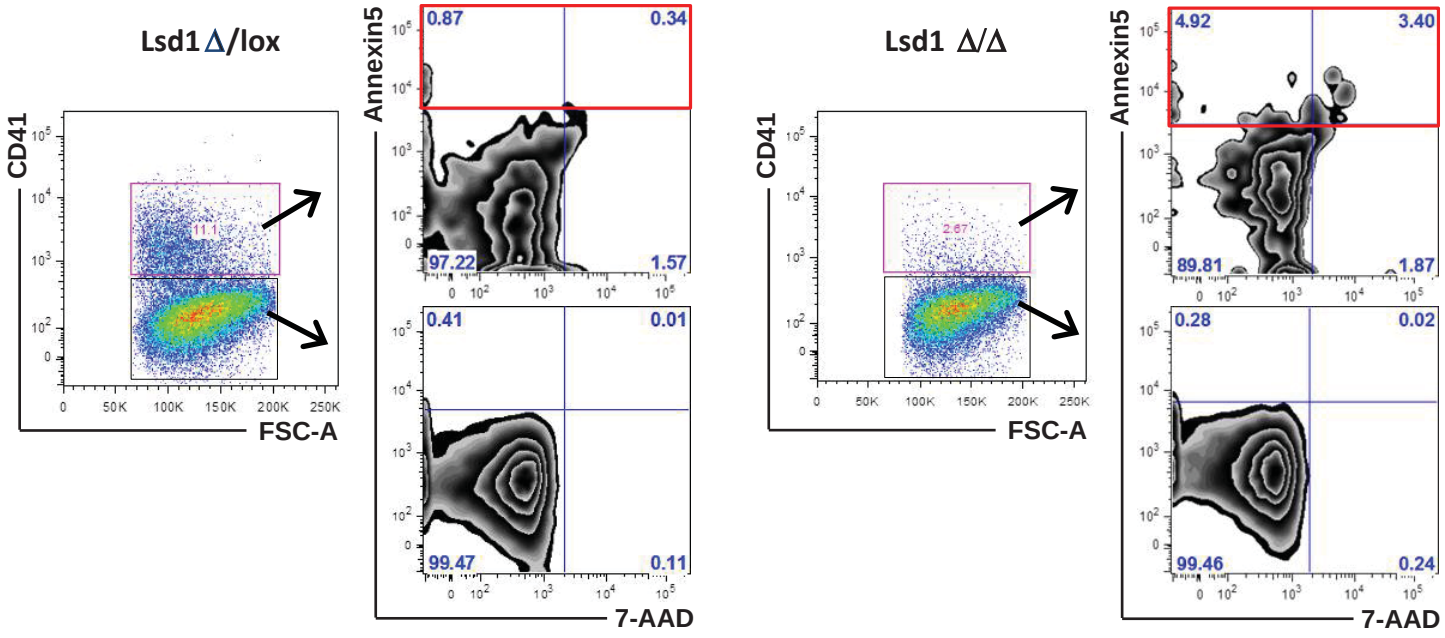
a


Day 1
Day 2
Day 3
CD41
CDH5
Lsd1Δ/lox
Lsd1Δ/Δ
2.39
1.07
93.1
3.41
11.6
8.88
69.5
10.1
30
8.4
58.3
3.28
2.47
1.18
93.6
2.78
2.36
8.4
77
12.3
3.42
12.1
80.8
3.79

b

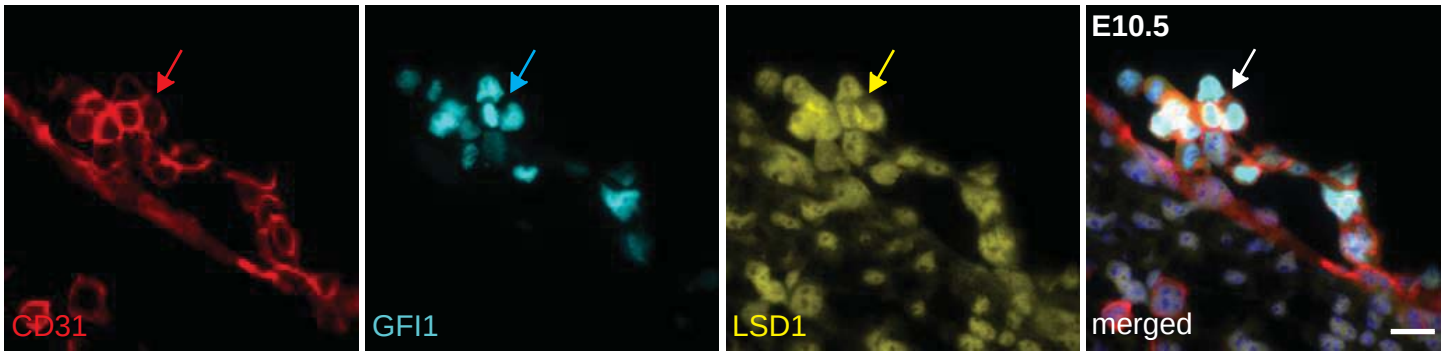

Lsd1 Δ/lox
Lsd1 Δ/Δ
EdU
DNA content
47.68
43.58
4.75
38.05
47.36
5.38
0
50K
100K
150K
200K
250K

c

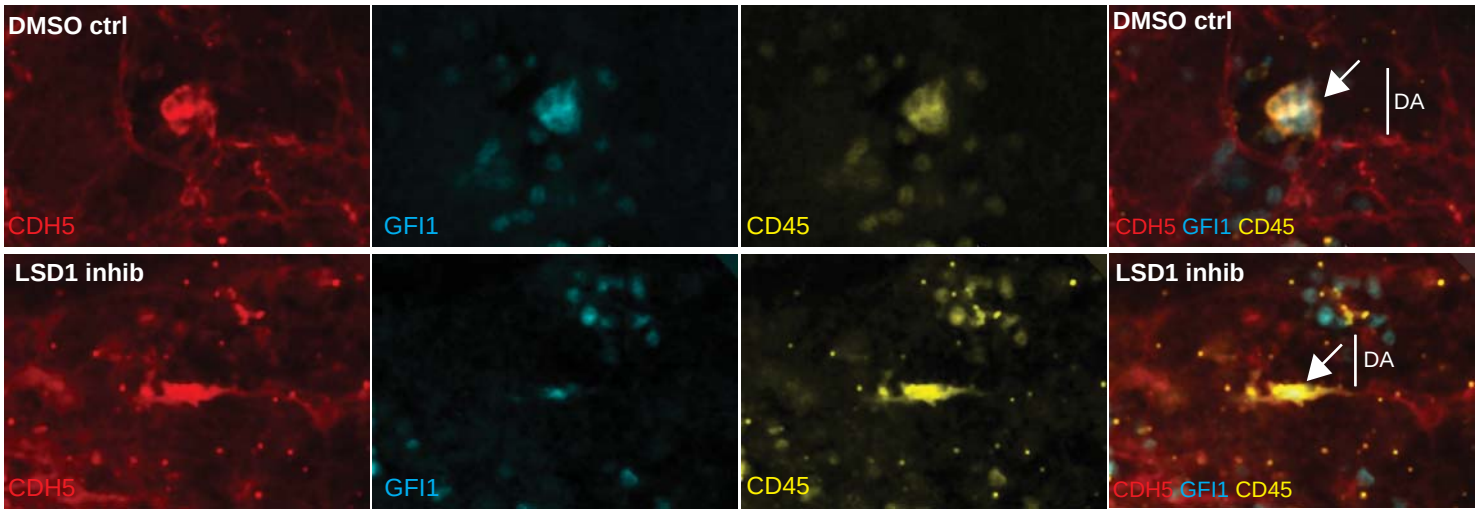

Lsd1 Δ/lox
Lsd1 Δ/Δ
CD41
FSC-A
Annexin5
7-AAD
11.1
0.87
0.34
97.22
1.57
0.41
0.01
99.47
0.11
2.67
4.92
3.40
89.81
1.87
0.28
0.02
99.46
0.24

S7

a

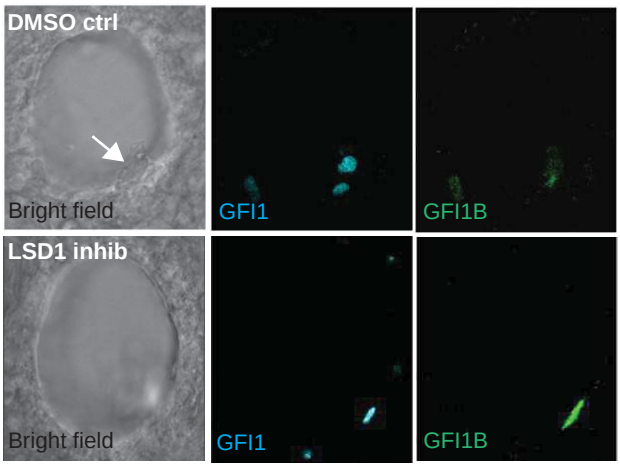


b



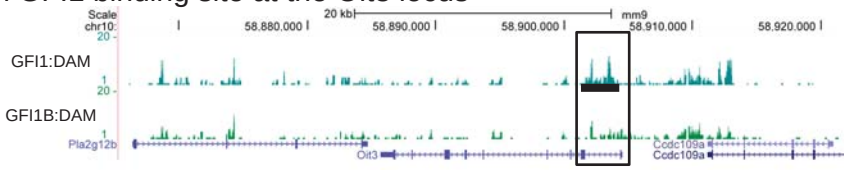
c

single channel images

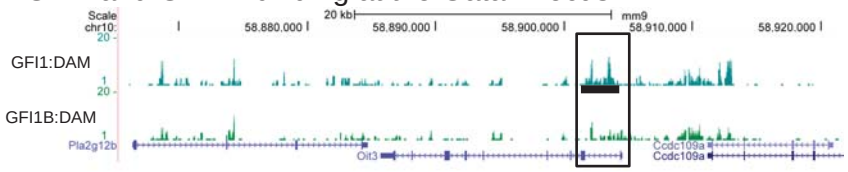


d

i GFI1 binding site at the Oit3 locus



ii GFI1 and GFI1B binding at the Gata-2 locus



iii GFI1B binding at the Hey-2 locus